Network-based multiple sclerosis pathway analysis with GWAS data from 15,000 cases and 30,000 controls

Baranzini, Sergio ; Khankhanian, Pouya ; Patsopoulos, Nikolaos ; Martin, Roland

Abstract: Multiple sclerosis (MS) is an inflammatory CNS disease with a substantial genetic component, originally mapped to only the human leukocyte antigen (HLA) region. In the last 5 years, a total of seven genome-wide association studies and one meta-analysis successfully identified 57 non-HLA susceptibility loci. Here, we merged nominal statistical evidence of association and physical evidence of interaction to conduct a protein-interaction-network-based pathway analysis (PINBPA) on two large genetic MS studies comprising a total of 15,317 cases and 29,529 controls. The distribution of nominally significant loci at the gene level matched the patterns of extended linkage disequilibrium in regions of interest. We found that products of genome-wide significantly associated genes are more likely to interact physically and belong to the same or related pathways. We next searched for subnetworks (modules) of genes (and their encoded proteins) enriched with nominally associated loci within each study and identified those modules in common between the two studies. We demonstrate that these modules are more likely to contain genes with bona fide susceptibility variants and, in addition, identify several high-confidence candidates (including BCL10, CD48, REL, TRAF3, and TEC). PINBPA is a powerful approach to gaining further insights into the biology of associated genes and to prioritizing candidates for subsequent genetic studies of complex traits.

DOI: https://doi.org/10.1016/j.ajhg.2013.04.019

AJHG The American Journal of Human Genetics

Network-based pathway analysis in multiple sclerosis with GWAS data from 15,000 cases and 30,000 controls

--Manuscript Draft--

Manuscript Number:	AJHG-D-13-00106R3
Full Title:	Network-based pathway analysis in multiple sclerosis with GWAS data from 15,000 cases and 30,000 controls
Article Type:	Article
Keywords:	multiple sclerosis; genetics; GWAS; pathway analysis; protein interaction networks; bioinformatics
Corresponding Author:	Sergio Baranzini, PhD UCSF San Francisco, CA UNITED STATES
First Author:	IMSGC IMSGC
Order of Authors:	IMSGC IMSGC
	Sergio Baranzini, PhD
	Pouya Khankhanian
	Nikolaos Patsopoulos
	Michael Li
	Chris Cotsapas
	Jim Stankovich
	Helle Bach Søndergaard
	Maria Ban
	Nadia Barizzone
	Laura Bergamaschi
	David Booth
	Dorothea Buck
	Paola Cavalla
	Elisabeth G Celius
	Giancarlo Comi
	Manuel Comabella
	Alastair Compston
	Isabelle Cournu-Rebeix
	Sandra D'alfonso
	Vincent Damotte
	Lennox Din
	Bénédicte Dubois
	Irina Elovaara
	Federica Esposito
	Bertrand Fontaine
	Andre Franke

	An Goris
	Christiane Graetz
	Franca Guerini
	Pierre-Antoine Gourraud
	Léna Guillot-Noel
	David Hafler
	Hakon Hakonarson
	Per Hall
	Anders Hamsten
	Hanne F Harbo
	Bernhard Hemmer
	Jan Hillert
	Anu Kemppinen
	Ingrid Kockum
	Keijo Koivisto
	Malin Larsson
	Mark Lathrop
	Maurizio Leone
	Christina M Lill
	Fabio Macciardi
	Roland Martin
	Vittorio Martinelli
	Filippo Martinelli-Boneschi
	Jacob L McCauley
	Kjell-Morten Myhr
	Paola Naldi
	Tomas Olsson
	Margaret A Pericak-Vance
	Annette Oturai
	Franco Perla
	Mauri Reunanen
	Janna Saarela
	Safa Saker-Delye
	Marco Salvetti
	Finn Sellebjerg
	Per Soelberg Sørensen
	Anne Spurkland
	Graeme Stewart
	Bruce Taylor
	Pentti Tienari
	Juliane Winkelmann

	WTCCC2 WTCCC2
	Frauke Zipp
	Adrian J Ivinson
	Jonathan L Haines
	Stephen Sawcer
	Philip DeJager
	Stephen L Hauser
	Jorge R Oksenberg
Abstract:	Multiple sclerosis (MS) is an inflammatory disease of the central nervous system with a substantial genetic component, originally mapped to only the HLA region. In the last 5 years, a total of 7 GWAS and one meta-analysis successfully identified 57 non-HLA susceptibility loci. Here we merged nominal statistical evidence of association with physical evidence of interaction to conduct a protein interaction network-based pathway analysis on two large genetic studies in MS comprising a total of 15,317 cases and 29,529 controls. The distribution of nominally significant loci at the gene-level matched the patterns of extended linkage disequilibrium in regions of interest. We found that products of genome-wide significantly associated genes are more likely to interact physically, and belong to the same or related pathways. We next searched for sub-networks (modules) of genes (proteins) enriched in nominally associated loci within each study and identified those modules in common between the two studies. We demonstrate that these modules are more likely to contain genes with bona-fide susceptibility variants and, in addition, identify several high-confidence candidates (including BCL10, CD48, REL, TRAF3 and TEC). Protein interaction network-based pathway analysis is a powerful approach to gain further insights into the biology of associated genes and to prioritize candidates for subsequent genetic studies of complex traits.

Network-based pathway analysis in multiple sclerosis with GWAS data from 15,000 cases and 30,000 controls

The International Multiple Sclerosis Genetics Consortium (IMSGC)[a].

Correspondence:

Sergio E. Baranzini, PhD

Department of Neurology. University of California, San Francisco.

675 Nelson Rising Lane, San Francisco, CA. 94158.

E-mail: Sergio.Baranzini@ucsf.edu

[a] The complete author list can be found in the Supplemental Materials.

ABSTRACT

Multiple sclerosis (MS) is an inflammatory disease of the central nervous system with a substantial genetic component, originally mapped to only the HLA region. In the last 5 years, a total of 7 GWAS and one meta-analysis successfully identified 57 non-HLA susceptibility loci. Here we merged nominal statistical evidence of association with physical evidence of interaction to conduct a protein interaction network-based pathway analysis on two large genetic studies in MS comprising a total of 15,317 cases and 29,529 controls. The distribution of nominally significant loci at the gene-level matched the patterns of extended linkage disequilibrium in regions of interest. We found that products of genome-wide significantly associated genes are more likely to interact physically, and belong to the same or related pathways. We next searched for sub-networks (modules) of genes (proteins) enriched in nominally associated loci within each study and identified those modules in common between the two studies. We demonstrate that these modules are more likely to contain genes with bona-fide susceptibility variants and, in addition, identify several high-confidence candidates (including *BCL10*, *CD48*, *REL*, *TRAF3* and *TEC*). Protein interaction network-based pathway analysis is a powerful approach to gain further insights into the biology of associated genes and to prioritize candidates for subsequent genetic studies of complex traits.

INTRODUCTION

Genome-wide association studies (GWAS) are a powerful approach to examine the genetic components of complex diseases. A commonly utilized strategy in the analysis of GWAS involves the evaluation of individual markers using a genome-wide significance cutoff P value of 5×10^{-8}, under the assumption of independence among markers. This approach minimizes false discoveries and has indeed enjoyed remarkable success, uncovering multiple variants associated with complex diseases and traits [1]. However, the very small fraction of both the heritable component and the population disease burden explained by the polymorphisms identified in most GWAS initiatives suggest that a sizable proportion of risk alleles are still being missed by this strategy [2; 3]. It is likely that alternative analysis approaches to GWAS data that focus on the combined effects of many loci, each making a small contribution to overall disease susceptibility, may reveal insights into the genetic basis of common chronic disease. An interesting study by the International Schizophrenia Consortium proved that by analyzing markers en-masse with a significance threshold as modest as 0.1, important information can be obtained from a well powered GWAS [4]. More recently, a similar approach was applied to multiple sclerosis (MS [MIM 126200]) using data from two independent GWAS and implicated thousands of markers with $p<0.2$, suggesting a clear polygenic model of disease susceptibility [5].

Furthermore, it is highly probable that univariate, single-locus analysis results contain informative trends that, when viewed in the contexts of genetic networks and fundamental molecular pathways, can expose aspects of the polygenic basis of disease susceptibility. A number of efforts to study biologically meaningful combinations of genes and markers have been reported, ranging from the simple computation of over-representation of associated loci

in gene ontology or KEGG pathways [6], up to more elaborated methods using gene sets enrichment [7 8 9; 10]. An advanced modification of these methods incorporates the use of protein interaction networks and searches for sub-networks (modules) enriched within the associated genes. This approach increases the prior probability of an association by merging statistical evidence of marker/gene association with physical evidence of interaction among those gene products (proteins). Several versions of this approach have been reported in multiple complex traits including autoimmune and neurological diseases [11-18] _ENREF_11.

MS is a common inflammatory disease of the central nervous system with a well-documented genetic component [19; 20]. Seven moderately powered but independent GWAS and one meta-analysis were reported between 2007 and 2011, altogether identifying 23 non-HLA associated loci [21-27]. Later, a meta-analysis (referred to as meta2.5 in this study) including most of these samples was carried out, and evidence of association for 2 additional loci was reported [28]. In collaboration with the Wellcome Trust Case-Control Consortium 2 (WTCCC2), the International MS Genetics Consortium (IMSGC) recently completed the largest GWAS in MS to date (referred to as WTCCC2 in this study) and raised the number of non-HLA genetic loci associated with this disease to 57 [29].

Despite this notable progress, our understanding of MS genetics remains incomplete. To further unravel the missing heritability in MS, we conducted a protein interaction network-based pathway analysis (PINBPA) of these two largely independent GWAS datasets in MS (meta 2.5 and WTCCC2), totaling more than 15,000 cases and almost 30,000 controls. We found that proteins encoded by genes harboring risk variants are more likely to interact, and take part of

the same or related pathways. Furthermore, additional susceptibility variants were identified through this approach.

SUBJECTS AND METHODS

Datasets and preprocessing

P-values for all tested SNPs (summary-level data) were collected for two MS datasets (WTCCC2 and meta2.5) and one each for type 1 diabetes (T1D [MIM 222100]), rheumatoid arthritis (RA [180300]), Crohn's disease (CD [MIM 266600]), coronary artery disease (CAD [MIM 611139]), hypertension (HT [MIM 145500]) and type 2 diabetes (T2D [MIM 125853]) (all from WTCCC1) [30] to be used as control. All datasets are composed of samples of European descent. The WTCCC2 MS dataset [29] consists of 9,772 cases and 17,376 controls analyzed with the Illumina Human 660-Quad and Illumina 1.2M platforms. The meta2.5 dataset is an imputation-based meta-analysis with 2,529,394 unique SNPs [28] and includes all previously published GWAS in MS (totaling 5,545 cases and 12,153 controls) with minimum overlap (less than 10%) of cases with WTCCC2. Thus, these two datasets are considered independent in the context of the present study. Supplementary Table 1 summarizes details for each study used in this work. In order to enrich for potentially functional variants, each dataset was filtered so as to keep only those SNPs which were non-synonymous and potentially deleterious (classified as either probably or possibly damaging by POLYPHEN2[31]) , or located in 5' or 3' untranslated regulatory regions (UTR), transcription factor binding sites (TFBS), or histone binding sites. To further reduce the number of redundant SNPs we eliminated those that were in close LD ($R^2 > 0.9$).

All data used in this manuscript was obtained following procedures in accordance with the ethical standards of the responsible committees on human experimentation (institutional and national) and proper informed consent was obtained.

Computing gene-wise p-values and association blocks

Since this study examines the functional relationships of genes and proteins, we needed to consider gene-level significance. To that end, we used VEGAS, a previously described method to convert individual SNP into gene-wise p-values [32]. VEGAS assigns SNPs to each of 17,787 autosomal genes according to positions on the UCSC Genome Browser hg18 assembly. In order to capture regulatory regions and SNPs in LD, gene boundaries are defined as 50 kb beyond the 5′ and 3′ UTRs of each gene. VEGAS takes into account linkage disequilibrium (LD) patterns between markers within a gene by using Monte-Carlo simulations from the multivariate normal distribution based on the LD structure of a set of reference individuals (the HapMap2 CEU population). In VEGAS, the number of simulations per gene is determined adaptively. In the first stage, 10^3 simulations are performed. If the resulting empirical p value is less than 0.1, 10^4 simulations are then performed. If the empirical p value from 10^4 simulations is less than 0.001, the program will perform 10^6 simulations. At each stage, the simulations are mutually exclusive. For computational reasons, if the empirical p value is 0, then no more simulations will be performed. An empirical p value of 0 from 10^6 simulations can be interpreted as $p < 10^{-6}$, which exceeds a Bonferroni-corrected threshold of $p < 2.8 \times 10^{-6}$ (≈0.05/17,787; this threshold is likely to be conservative given the overlap between genes).

We defined association blocks as those groups of sequential genes with p-value < 0.05. A block_id was assigned to each association block along the genome for each study.

Protein interaction network-based pathway analysis (PINBPA)

We downloaded the entire iRefIndex database, a collection of 15 human protein interaction sets from different sources and computed the union dataset. This set comprised more than 400,000 interactions among ~25,000 proteins. However, many of these interactions were either predicted or backed up by a single experiment (i.e. a single publication). In order to minimize the rate of false positives, we then filtered this large network to keep only those interactions that were described in at least two independent publications. This resulted in a network of 8,960 proteins (nodes) and 27,724 interactions (edges). We used this high-confidence network for all subsequent analyses. The network was uploaded into Cytoscape 2.8.2 and annotated with genomic position, and gene-wise p-value, block_id and bona-fide genes (loaded as node attributes) for all studies analyzed. To avoid the complexity of the HLA region, p-values for genes mapping to the 6p21.3 region were not included as attributes. However, the nodes corresponding to those genes were left in the network, as they may still participate in relevant sub-networks with other significant genes.

For each dataset, significant first-order interactions were computed by filtering the main network so as to keep only those genes (proteins) with VEGAS p-values <0.05. Then, the number of resulting nodes, edges and size of the largest connected component were computed within Cytoscape. To evaluate the likelihood that these numbers were obtained by chance (as a

consequence of the sheer number of interactions), we computed 1,000 simulations assigning p-values at random from the same network and creating sub-networks of similar size. These simulations were used as background to estimate the significance of the sub-networks obtained with the real gene-wise p-values.

We then used the program (plugin) *jActiveModules* to conduct searches of sub-networks that are enriched in (but not necessarily composed of) genes with significant p-values. Although *jActiveModules* was originally designed to discover "active" sub-networks by evaluating network connectedness among differentially expressed transcripts, we adapted it to take association p-values instead [33]. *jActiveModules* starts by converting each gene p-value into a z-score by using the inverse normal cumulative distribution function (CDF). Then an aggregate z-score is produced (z_A) for an entire sub-network A of *k* genes, by summing the z_i over all genes in the sub-network $Z_A = \frac{1}{\sqrt{k}} \sum_{i \in A} Z_i$.

In order to properly capture the connection between genetic association and network topology, the probability of obtaining a given z_A score by chance must be evaluated. This is accomplished by randomly samples gene sets of size k using a Monte Carlo approach, computing their z_A scores, and then use these to derive estimates for the score mean μ_k and standard deviation σ_k for each *k*. Because the means and standard deviation are expected to be a smooth function of k, noise in the Monte Carlo estimates could be reduced by using a sliding window average. Thus the corrected sub-network score SA is $S_A = \frac{(Z_A - \mu_k)}{\sigma_k}$. We took an S_A>3 as evidence of a biologically active sub-network.

Gene ontology and cell-specific expression of candidate genes

To evaluate the biological relevance of the 79 genes described in Table 2, we conducted a gene ontology analysis (Biological process FAT set) using DAVID [34] with the following parameters: Similarity term overlap = 10, Similarity threshold = 0.50, Initial group membership = 5, Final group membership = 3, Multiple linkage threshold = 0.50 and EASE = 0.01. For pathways, we used the KEGG set and default parameters.

Cell-specific expression was assessed using the Gene enrichment profiler tool[35]. This tool computes the expression and enrichment of any set of query genes based on a reference set obtained from 126 normal tissues and cell types (represented by 557 microarrays).

Additional analysis and plots were performed using the R statistical package.

Domain knowledge score (DKS)

To prioritize unreported associations, we used a custom tool named domain knowledge score (DKS). DKS was programmed in R and works by performing sequential automated PubMed searches with each gene from a custom list and any combination of search terms. In this article, we combined each gene symbol of interest with the terms 'multiple sclerosis' OR 'inflammation' OR 'immunity'. In order to also capture older articles that may refer to outdated gene identifications, the tool also searches for all synonyms and aliases within a specific species. The score each gene gets is simply the number of PubMed articles (excluding reviews) retrieved with the input search terms. The DKS tool is available upon request.

RESULTS

Here we describe a multi-analytical approach to integrate two large genomic datasets in MS (Supplementary Figure 1). Through this approach we merge statistical evidence of association with physical evidence of interaction at the protein level to identify associated loci and highlight functional pathways involved in disease susceptibility.

Nominal gene-level associations cluster into blocks

Individual SNP-wise summary-level data from two largely independent GWAS in MS was utilized to compute gene-level p-values using VEGAS [32]. The first of these two studies comprised 9,772 cases and 17,376 controls and was recently published by the International MS genetics consortium (IMSGC) and the Wellcome Trust Case Control Consortium (WTCCC2) 2 [29]. The second study is a meta-analysis encompassing all previous GWAS studies in MS, including a total of 5,545 cases and 12,153 controls [28]. In order to maximize the chance that variants had a functional impact on the encoded protein, we selected the subset of 137,457 SNPs which were non-synonymous and potentially deleterious (classified as either probably or possibly damaging by POLYPHEN2), or located in 5' or 3' untranslated regulatory regions (UTR), transcription factor binding sites (TFBS), or histone binding sites. VEGAS computes gene-wise p-values by taking into account relative genomic position, number of SNPs within a gene and patterns of LD for the appropriate ethnic background. An adaptive simulation strategy is used to calculate an empirical gene-based p-value for each annotated gene, which defines $p < 2.8 \times 10^{-6}$ as Bonferroni significant. Because our main hypothesis states that even modestly associated genes may participate in biologically plausible pathways, we considered all genes with VEGAS $p < 0.05$. A Manhattan plot visualization of both datasets at the gene-level denotes the presence of

peaks of association, similar to those observed with SNP-level data (Figure 1). The distribution of nominally significant loci at the gene-level largely replicated between the studies (see example in gray box in Figure 1) and closely matched the patterns of extended linkage disequilibrium previously observed in regions of interest. Specifically, 665 association blocks containing 1,997 genes were identified for the WTCCC2 dataset, and 612 blocks containing 1,707 genes for the meta2.5 dataset. Of these, 625 genes overlapped, representing a much higher than expected proportion (4.8-fold enrichment) compared to what would be expected by chance (χ^2 test, $p < 2.2\ 10^{-16}$). Notably, association blocks defined in this way closely match the boundaries of the association regions for the 57 MS associated loci recently reported by the IMSGC [29], defined by extending a fixed genetic distance (0.25cM) from the lead SNP and from there to the closest recombination hot-spot from HapMap2. The overlap between studies was still significant after excluding genes from the MHC and from blocks implicated by the 57 WTCCC2 loci and the 2 additional meta2.5 associated SNPs. In this filtered set, we found nominal association in 557 blocks (1471 genes) from WTCCC2 and 530 (1298 genes) from meta2.5 with an overlap of 271 genes (3.2-fold enrichment, Fisher's exact $p < 2.2\ 10^{-16}$).

Protein network-based pathway analysis

We next sought to identify additional MS susceptibility loci by combining statistical evidence of gene association with physical evidence of interaction of their respective gene products using a curated human protein interaction network (PIN) dataset consisting of 8,960 proteins (nodes) and 27,724 interactions (edges) (see Methods). All subsequent experiments were performed

using Cytoscape, an open-source, and extensible network visualization and analysis tool [36]. When we extracted the nodes with p-values < 0.05, sub-networks of 838 nodes (401 edges) and 761 nodes (304 edges) were generated for the WTCCC2 and meta2.5 datasets, respectively (we refer to these as first-order networks). Given that neighboring genes have been shown to be functionally related and thus more likely to interact [18; 37], we repeated this experiment ensuring that only one gene per block was extracted from the main network. This resulted in first-order networks of 462 nodes (183 edges) for WTCCC2 and of 414 nodes (147 edges) for the meta2.5 dataset. Both sub-networks were more connected than would be expected by chance, as demonstrated by a simulation experiment in which 1,000 networks of similar size were extracted from the same PIN at random (Figure 2). Of the other datasets used as controls T1D, CD and RA also produced highly connected sub-networks (Figure 2A). In each case, sub-networks were composed of a large connected component and several smaller networks or isolated nodes (singletons). When first-order networks were computed using more significant p-value thresholds, most diseases showed more connections than expected (Supplementary Figure 2). We also tested whether the size of the main component was higher than what it would be expected by chance (given the number of edges in the first-order network) and observed that less than 1% of random networks resulted in larger connected components than those obtained for WTCCC2, meta 2.5 and CD datasets. Approximately 10% of random networks resulted in connected components of the size of those generated by HT and RA (Figure 2B). Again, when first-order networks were computed using more significant p-value thresholds, most diseases showed larger connected components than expected (Supplementary Figure 3). The higher than expected first-order interactions and size of the

main connected component of these networks suggests a biologically plausible mechanism by which these gene sets coordinately affect cellular behavior.

Given the small-world topology of the human protein interactome, it is possible that a few highly connected nodes (hubs) bring together several associated genes, even though the hubs themselves are not associated, thus defining biologically associated modules. To explore this possibility, we conducted searches for sub-networks enriched in significant genes using *jActive modules*, a Cytoscape plugin based on a greedy heuristic algorithm with internal cross-validation [33]. Fifteen significant and minimally overlapping modules of sizes 5-200 were identified for the WTCCC2 dataset. Similarly, 16 significant and minimally overlapping modules of sizes 5-189 were identified for the meta2.5 dataset (Table 1). We next computed the union of all modules within each dataset, resulting in a single connected network of 464 nodes and 820 edges for WTCCC2 and another of 605 nodes and 1,031 edges for meta2.5. Finally, we computed the intersection of these two networks, which yielded 118 nodes and 95 edges. Of these, 88 genes were arranged in 13 networks of sizes 2-27, while the remaining 30 genes remained as singletons (Figure 3). We concentrated on the 88 genes arranged in networks as these genes and the connections among them were independently identified in both MS studies, and as such, we hypothesized these would have higher potential to include bona-fide susceptibility loci. Of these 88 genes, 54 had nominally significant p-values in both WTCCC2 and meta2.5 studies (v-shaped nodes in Figure 3), while the remaining 34 had significant p-values in only one or neither study. These 54 genes are of highest importance to our approach since they have significant p-values in both studies, and they were identified as components of significant networks in both studies as well. Notably, 30 of these genes contain either bona-fide

susceptibility variants (n=13), or are located within bone-fide associated regions (n=17) (Table 2, Figure 4), thus representing a specificity of 56% (Supplementary Table 2 lists all blocks harboring genes with bona-fide susceptibility variants and allows comparison of block structure from WTCCC2 and meta 2.5). Considering that only 34 of the 57 MS susceptibility loci identified to date are represented in the PIN, this approach was able to identify bona-fide MS loci with a sensitivity of 88%. Although an independent replication is warranted to firmly establish whether the remaining 24 genes are indeed associated, the high recall observed using the network-based approach lends support to their involvement in MS susceptibility (Supplementary Table 3 lists the complete block structure of each of these candidates in both WTCCC2 and meta 2.5 studies).

When we explored the 30 singleton genes in the intersection network, we found that 26 of them had nominally significant p-values in both studies (Table 2, bold entries). Although these genes ended up as singletons in the intersection network, each of them was part of a connected network in either the WTCCC2 or meta 2.5 individual studies. Therefore, we also evaluated to what extent these genes (significant in both studies, but participating in networks in only one of them), included bona-fide MS susceptibility loci. Of these 26 genes, 12 either contain bona-fide susceptibility variants (n=3) or are located within bone-fide associated regions (n=9), representing a specificity of 46% and a sensitivity of 35% (Table 1, Figure 4).

As a control, we also evaluated the recall potential of the 154 genes with nominally significant p-values in both studies, but that were not found in networks. Only 13 of them (8%) were bona-fide MS genes and 26 (17%) were located within bona-fide MS blocks. The remaining 115 (75%)

remained potential(non-validated) associations (Figure 4). These findings represent a sensitivity of 68%, and a specificity of 25%.

Altogether, these results suggest that even nominally significant genes, if replicated in more than one study, represent a select list of candidates for further analysis. However, in the absence of any additional evidence the chances of discovering genuine associations among these genes are still hampered by a significant proportion of false positives. These probabilities are significantly increased when considering those genes that, in addition to showing (nominally) significant associations also participate in interaction networks in at least one study. The best results, nonetheless, were obtained when nominally significant genes were also identified as part of the same interaction network in both studies.

The analysis of other diseases from WTCCC1 used as control also support this interpretation. In those datasets, the average sensitivity was 42.3% and the average specificity was 8.3%. Notably, CD yielded a sensitivity of 88% (the same obtained for MS), albeit with a much more modest specificity of 16% (compared with 56% for MS). The main factors contributing to the significantly better performance of the MS datasets were most likely their size and the availability of a replication dataset.

Biological significance of associated and candidate genes in MS

To explore the biological significance of the genes with either confirmed or suspected role in MS susceptibility (described in Table 2) we conducted a gene ontology (GO) and pathway analysis using DAVID. Among the 79 genes in these lists, GO analysis (Biological process)

identified three main categories as significantly enriched: leukocyte activation (enrichment score=9, lead category FDR-corrected p-value=1.3 x 10^{-8}), apoptosis (enrichment score=6.16, lead category FDR-corrected p-value=2.2 x 10^{-6}) and positive regulation of macromolecule metabolic process (enrichment score=5.86, lead category FDR-corrected p-value=4.7 x 10^{-8}). When KEGG pathways were evaluated, JAK-STAT signaling pathway (enrichment score=3.47, lead category FDR-corrected p-value=1.4 x 10^{-5}), Acute myeloid leukemia (enrichment score=2.22, lead category FDR-corrected p-value=5.9 x 10^{-3}), and T cell receptor signaling (enrichment score=1.63, lead category FDR-corrected p-value=0.01) were significantly enriched.

We also computed the tissue specificity of these genes using the gene enrichment profiler utility (see methods). Approximately two thirds of these genes were highly expressed in immune related cell types, and about half in the CNS (Supplementary Figure 4, red or black color). However, enrichment maps overwhelmingly highlight immune related cell types (Supplementary Figure 5). Since only three genes (*PDE4A [MIM 600126]*, *RAB3A* [MIM 179490] and *VAMP1 [MIM 185880]*) appear to be specifically enriched in the CNS, we were unable to confirm our earlier observation that neural pathways were involved in the susceptibility to MS[12].

Finally, we used a combination of gene-level statistical significance and text mining (DKS, domain knowledge score, see Methods) to highlight some of the previously unreported candidate associations emerging from the analysis (Table 2). Based on stringent criteria ($p<0.01$ in both studies and DKS >50) 5 genes were identified as the most plausible candidates. B-cell lymphoma 10 (*BCL10* [MIM 603517]) (DKS=62), *CD48* [MIM 109530] (also known as B cell

membrane protein) (DKS=83), v-rel reticuloendotheliosis viral oncogene homolog (*REL* [MIM 164910]) (DKS=630), TNF receptor-associated factor 3 (*TRAF3* [MIM 601896]) (DKS=60), and the protein tyrosine kinase *TEC* [MIM 600583] (DKS=230). Although it is not possible to unequivocally implicate any of these candidates in MS susceptibility, in the absence of experimental functional data the combined strategy described here provides a more comprehensive interpretation of these associations.

DISCUSSION

One plausible cause for the manifestation of complex diseases is the genetic alteration in the function of specific biological pathways through the presence of multiple variants in different genes (each contributing a modest amount to disease predisposition) and ultimately leading to disruptions in normal biological processes. We have found that even nominally associated genes (i.e. gene-level data) were not scattered randomly across the genome, but rather agglomerated into clusters or blocks of association, in a similar fashion that regional association plots show for SNP-level data. In fact, the gene-wise association blocks defined in this study and the critical regions defined in the original WTCCC2 publication are remarkably similar (see supplementary Table 2). It is noteworthy that any other gene-wise p-value threshold would have resulted in a different arrangement of genes into blocks, likely smaller and fewer. Thus, the close agreement in association block structure and size supports our choice of nominal p-value as a threshold for the remainder of the study. Furthermore, this finding has important implications, as it indicates that our strategy of selecting potentially functional SNPs and nominally significant genes produces comparable results to the more established approach utilized in our previous study of extending a fixed genetic distance from the lead SNP, and from there to the next recombination hot-spot [29]. This also suggests that in most regions, the extended patterns of LD would determine the upper limit of resolution of this approach, except in cases in which a variant with obvious functional consequence is identified within these regions.

We demonstrated that truly associated genes are more likely to be connected in the PIN. By extension, we hypothesized that significant sub-networks (enriched in nominally significant

genes) would contain genes which are more likely to be genuinely associated. Assuming that 10^7 common single nucleotide variants exist in the human genome and that 100 of them are truly associated with MS, the prior probability of finding an association by chance is 100,000:1 against (10^{-5}). Theoretical calculations have suggested that the cutoff of statistical significance required to yield an association that is more likely true than false, is directly related to its sample size (power) [38]. For example, under these assumptions, a p-value of 10^{-6} is predicted to identify an association that is 10 times more likely to be true than false for a study of 10,000 cases and 10,000 controls, but equally likely to be true or false if the size of the study is 2,000 cases and 2,000 controls. For a study with 1,000 cases and 1,000 controls, that same threshold p-value will identify associations which are 10 times more likely to be false than true. These theoretical estimates have also showed that if the prior probability of an association is increased, for example, by two orders of magnitude (from 10^{-5} to 10^{-3}), the p-value threshold generating the same level of confidence in a result can be reduced by roughly the same magnitude (from 10^{-6} to 10^{-4}). It follows that increasing the prior probability is a meaningful way to increase the power of detecting bona-fide associations in a study of a given size. Several ways to increase the prior probability of an association exist. In this study we have aimed at increasing the prior odds by using a 3-way strategy. First, we conducted our analysis only using functional or potentially functional SNPs. Since non-synonymous coding variants and variants in regulatory regions or splice sites are more likely to have a functional effect than variants in silent non-coding regions concentrating analysis on these more functional relevant variants is a reasonable strategy to improve the prior odds [39]. Second we computed gene-wise p-values, thus significantly reducing the number of possible tests by ~8-fold (from 137,457 to 17,787).

Third we prioritized candidates that were arranged in interaction networks, which as shown above, increases the likelihood of finding true associations.

Altogether, this strategy (similar in concept to the genomic convergence paradigm previously described [40]) is likely to increase the priors, although it is not possible to determine exactly by which magnitude. The fact that recall power of the two MS datasets was large (AUC of 0.95 and 0.88), further encouraged us to evaluate genes with an even modest statistical evidence of association. Therefore, we took a liberal approach and considered every gene-level association with a nominal p-value of 0.05. In support of this strategy, we found that the overlap of nominally significant genes between WTCCC2 (n=1,997) and meta2.5 (n=1,707) studies was 625, a 6-fold increase of what it would be expected by chance (Fisher's exact p-value $<10^{-16}$).

An important finding of this study is that nominally associated genes are more connected in the PIN than what it would be expected by chance. This provides further evidence that in well powered studies, the 3-way strategy followed here (selection of potentially functional SNPs, analyzing nominal gene-level significance and studying genes in the context of biological networks) maximizes the potential of finding bona-fide associations. Furthermore, this approach may highlight the importance of a different gene than the one originally selected within a GWAS associated block. For example, a non-synonymous SNP (rs3748816) in membrane metalloendopeptidase-like 1 (*MMEL1*) was originally identified as a susceptibility gene though a screen of candidate genes in 3,444 affecteds and 2,595 controls showing a p-value of 3.54×10^{-6} (odds ratio=1.16) [41]. This association was further replicated by the WTCCC2 GWAS with a p-value of $2.25\ 10^{-13}$. Its p-value in meta 2.5 was $2.81\ 10^{-5}$. However, given the

extensive LD in this region, it is not possible to exclude the possibility that other genes within this block are instead associated with MS. In addition to *MMEL1*, genes in this association block include *PLCH2 [MIM612836]*, *PANK4 [MIM 606162]*, *HES5 [MIM 607348]*, *TNFRSF14 [MIM* 602746] and *C1orf93*. This region was also identified in the present study with a block p-value of 10^{-7} (genome-wide significant) in WTCCC2 and 1.49 10^{-4} in meta2.5 (Supplementary Table 2). However, the only gene product from this block that appeared in the final intersection network was TNFRSF14, with direct interactions with TRAF2 (not significant) and TRAF3 (p-values of 1.3 10^{-3} and 1.13 10^{-3} in WTCCC2 and meta2.5 respectively)(Figure 4). Furthermore, TNFRSF14 is a ligand of TNFSF14, one of the 57 susceptibility loci identified by the WTCCC2 GWAS. Interestingly, although physically within *MMEL1*, rs3748816 was mapped to *TNFRSF14* by VEGAS in WTCCC2 and to both genes in meta2.5, likely due to its high LD. Altogether, and in the absence of additional functional experimental data, these results provide more evidence to implicate variants in *TNFRSF14* than *MMEL1* as an MS susceptibility locus. Ultimately, however, experimental evidence will be needed to determine this with precision.

Another example is the association block containing *VCAM1 [MIM 192225]*, *EXTL2 [MIM 602411]*, *SLC30A7 [MIM 611149]*, *DPH5 [MIM 611075]*, and *S1PR1 [MIM 601974]*. Although VCAM has been selected as the most likely associated gene from this block (presumably due to its function in cell adhesion), the WTCCC2 regional association plot shows that this gene falls slightly outside the block, and the most significant functional SNP maps to *SLC30A7*. In this study, however, the block extends to include *VCAM1* and *S1PR1*. Furthermore, the only gene that is significant in both WTCCC2 and meta2.5 and shows in the final intersection network is *S1PR1* (sphingosine-1-phosphate receptor 1) (Figure 4). This is of relevance, as *S1PR1* is the

target of the disease modifying therapy Fingolimod. Again, further experimental approaches are warranted to determine which are the functionally relevant associations in each of these loci.

We acknowledge that the lack of an independent replication is a limitation when predicting new associations. However, the successful identification of several bona-fide susceptibility variant-containing genes, the prioritization of different genes within a known association block and the proposal of new candidate associations are valuable outcomes only achieved by integration of different sources of evidence. Results from this approach contribute to firmly establish that genes and pathways involved in the immune response are the major drivers of MS risk.

Supplemental Data description:

Supplementary Figure 1. Strategy

Supplementary Figure 2. Total number of edges as a function of the number of significant number of genes at different significance thresholds

Supplementary Figure 3. Size of the largest connected component as a function of the total number of edges in the first-order networks at different thresholds of significance

Supplementary Figure 4. Transcript expression heatmap.

Supplementary Figure 5. Transcript enrichment heatmap

Supplementary Table 1. Study details

Supplementary Table 2. Prioritization of candidate genes within a block (Functional mapping)

Supplementary Table 3. Previously unreported candidate genes

Acknowledgements

We thank the research subjects, their families and the healthy control individuals that participated in this study. SEB and PDJ are Harry Weaver Neuroscience Scholars from the US National MS Society. BD is a Clinical Investigator of the Research Foundation Flanders (FWO-Vlaanderen) and AG and BD are supported by the Research Fund KU Leuven (OT/11/087).

None of the authors declared any conflict of interest.

Web Resources

Online Mendelian Inheritance in Man (http://www.omim.org)

Gene expression profiler tool: http://xavierlab2.mgh.harvard.edu/EnrichmentProfiler/

References

1. Hindorff, L.A., Sethupathy, P., Junkins, H.A., Ramos, E.M., Mehta, J.P., Collins, F.S., and Manolio, T.A. (2009). Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. Proc Natl Acad Sci U S A 106, 9362-9367.
2. Bodmer, W., and Bonilla, C. (2008). Common and rare variants in multifactorial susceptibility to common diseases. Nature genetics 40, 695-701.
3. Eberle, M.A., Ng, P.C., Kuhn, K., Zhou, L., Peiffer, D.A., Galver, L., Viaud-Martinez, K.A., Lawley, C.T., Gunderson, K.L., Shen, R., et al. (2007). Power to detect risk alleles using genome-wide tag SNP panels. PLoS genetics 3, 1827-1837.
4. Purcell, S.M., Wray, N.R., Stone, J.L., Visscher, P.M., O'Donovan, M.C., Sullivan, P.F., and Sklar, P. (2009). Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. Nature 460, 748-752.
5. (IMSGC), I.M.S.G.C., Bush, W.S., Sawcer, S.J., de Jager, P.L., Oksenberg, J.R., McCauley, J.L., Pericak-Vance, M.A., and Haines, J.L. (2010). Evidence for polygenic susceptibility to multiple sclerosis--the shape of things to come. American journal of human genetics 86, 621-625.
6. Torkamani, A., Topol, E.J., and Schork, N.J. (2008). Pathway analysis of seven common diseases assessed by genome-wide association. Genomics 92, 265-272.
7. Wang, L., Jia, P., Wolfinger, R.D., Chen, X., and Zhao, Z. (2011). Gene set analysis of genome-wide association studies: methodological issues and perspectives. Genomics 98, 1-8.
8. Wang, K., Li, M., and Bucan, M. (2007). Pathway-Based Approaches for Analysis of Genomewide Association Studies. Am J Hum Genet 81, 1278-1283.
9. Holmans, P., Green, E.K., Pahwa, J.S., Ferreira, M.A., Purcell, S.M., Sklar, P., Owen, M.J., O'Donovan, M.C., and Craddock, N. (2009). Gene ontology analysis of GWA study data sets provides insights into the biology of bipolar disorder. American journal of human genetics 85, 13-24.
10. Chen, L.S., Hutter, C.M., Potter, J.D., Liu, Y., Prentice, R.L., Peters, U., and Hsu, L. (2010). Insights into colon cancer etiology via a regularized approach to gene set analysis of GWAS data. American journal of human genetics 86, 860-871.
11. Lee, I., Blom, U.M., Wang, P.I., Shim, J.E., and Marcotte, E.M. (2011). Prioritizing candidate disease genes by network-based boosting of genome-wide association data. Genome research 21, 1109-1121.
12. Baranzini, S.E., Galwey, N.W., Wang, J., Khankhanian, P., Lindberg, R., Pelletier, D., Wu, W., Uitdehaag, B.M., Kappos, L., Polman, C.H., et al. (2009). Pathway and network-based analysis of genome-wide association studies in multiple sclerosis. Hum Mol Genet 18, 2078-2090.
13. Jia, P., and Zhao, Z. (2012). Searching joint association signals in CATIE schizophrenia genome-wide association studies through a refined integrative network approach. BMC Genomics 13 Suppl 6, S15.
14. Fernandez, R.M., Bleda, M., Nunez-Torres, R., Medina, I., Luzon-Toro, B., Garcia-Alonso, L., Torroglosa, A., Marba, M., Enguix-Riego, M.V., Montaner, D., et al. (2012). Four new loci associations discovered by pathway-based and network analyses of the genome-wide variability profile of Hirschsprung's disease. Orphanet J Rare Dis 7, 103.

15. Lee, Y.H., Bae, S.C., Choi, S.J., Ji, J.D., and Song, G.G. (2012). Genome-wide pathway analysis of genome-wide association studies on systemic lupus erythematosus and rheumatoid arthritis. Mol Biol Rep 39, 10627-10635.
16. Jensen, M.K., Pers, T.H., Dworzynski, P., Girman, C.J., Brunak, S., and Rimm, E.B. (2011). Protein interaction-based genome-wide analysis of incident coronary heart disease. Circulation Cardiovascular genetics 4, 549-556.
17. Jia, P., Zheng, S., Long, J., Zheng, W., and Zhao, Z. (2011). dmGWAS: dense module searching for genome-wide association studies in protein-protein interaction networks. Bioinformatics 27, 95-102.
18. Rossin, E.J., Lage, K., Raychaudhuri, S., Xavier, R.J., Tatar, D., Benita, Y., Cotsapas, C., and Daly, M.J. (2011). Proteins encoded in genomic regions associated with immune-mediated disease physically interact and suggest underlying biology. PLoS genetics 7, e1001273.
19. Compston, A. (2005). McAlpine's multiple sclerosis.(Philadelphia: Churchill Livingstone Elsevier).
20. Hauser, S.L., and Goodin, D.S. (2012). Multiple Sclerosis and other demyelinating diseases. In Harrison's principles of internal medicine, D.I. Longo, ed. (New York, McGraw-Hill), pp 3395-3409.
21. (IMSGC), I.M.S.G.C., Hafler, D.A., Compston, A., Sawcer, S., Lander, E.S., Daly, M.J., De Jager, P.L., de Bakker, P.I., Gabriel, S.B., Mirel, D.B., et al. (2007). Risk alleles for multiple sclerosis identified by a genomewide study. N Engl J Med 357, 851-862.
22. Baranzini, S.E., Wang, J., Gibson, R.A., Galwey, N., Naegelin, Y., Barkhof, F., Radue, E.W., Lindberg, R.L., Uitdehaag, B.M., Johnson, M.R., et al. (2009). Genome-wide association analysis of susceptibility and clinical phenotype in multiple sclerosis. Human molecular genetics 18, 767-778.
23. ANZgene. (2009). Genome-wide association study identifies new multiple sclerosis susceptibility loci on chromosomes 12 and 20. Nature genetics 41, 824-828.
24. Jakkula, E., Leppa, V., Sulonen, A.M., Varilo, T., Kallio, S., Kemppinen, A., Purcell, S., Koivisto, K., Tienari, P., Sumelahti, M.L., et al. (2010). Genome-wide association study in a high-risk isolate for multiple sclerosis reveals associated variants in STAT3 gene. American journal of human genetics 86, 285-291.
25. Sanna, S., Pitzalis, M., Zoledziewska, M., Zara, I., Sidore, C., Murru, R., Whalen, M.B., Busonero, F., Maschio, A., Costa, G., et al. (2010). Variants within the immunoregulatory CBLB gene are associated with multiple sclerosis. Nature genetics 42, 495-497.
26. Nischwitz, S., Cepok, S., Kroner, A., Wolf, C., Knop, M., Muller-Sarnowski, F., Pfister, H., Roeske, D., Rieckmann, P., Hemmer, B., et al. (2010). Evidence for VAV2 and ZNF433 as susceptibility genes for multiple sclerosis. Journal of neuroimmunology 227, 162-166.
27. Comabella, M., Craig, D.W., Camina-Tato, M., Morcillo, C., Lopez, C., Navarro, A., Rio, J., Montalban, X., and Martin, R. (2008). Identification of a novel risk locus for multiple sclerosis at 13q31.3 by a pooled genome-wide scan of 500,000 single nucleotide polymorphisms. PLoS ONE 3, e3490.
28. Patsopoulos, N.A., Esposito, F., Reischl, J., Lehr, S., Bauer, D., Heubach, J., Sandbrink, R., Pohl, C., Edan, G., Kappos, L., et al. (2011). Genome-wide meta-analysis identifies novel multiple sclerosis susceptibility loci. Annals of neurology 70, 897-912.

29. (IMSGC), I.M.S.G.C., Sawcer, S., Hellenthal, G., Pirinen, M., Spencer, C.C., Patsopoulos, N.A., Moutsianas, L., Dilthey, A., Su, Z., Freeman, C., et al. (2011). Genetic risk and a primary role for cell-mediated immune mechanisms in multiple sclerosis. Nature 476, 214-219.
30. WTCCC. (2007). Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature 447, 661-678.
31. Adzhubei, I.A., Schmidt, S., Peshkin, L., Ramensky, V.E., Gerasimova, A., Bork, P., Kondrashov, A.S., and Sunyaev, S.R. (2010). A method and server for predicting damaging missense mutations. Nature methods 7, 248-249.
32. Liu, J.Z., McRae, A.F., Nyholt, D.R., Medland, S.E., Wray, N.R., Brown, K.M., Hayward, N.K., Montgomery, G.W., Visscher, P.M., Martin, N.G., et al. (2010). A versatile gene-based test for genome-wide association studies. American journal of human genetics 87, 139-145.
33. Ideker, T., Ozier, O., Schwikowski, B., and Siegel, A.F. (2002). Discovering regulatory and signalling circuits in molecular interaction networks. Bioinformatics 18 Suppl 1, S233-240.
34. Dennis, G., Jr., Sherman, B.T., Hosack, D.A., Yang, J., Gao, W., Lane, H.C., and Lempicki, R.A. (2003). DAVID: Database for Annotation, Visualization, and Integrated Discovery. Genome biology 4, P3.
35. Benita, Y., Cao, Z., Giallourakis, C., Li, C., Gardet, A., and Xavier, R.J. (2010). Gene enrichment profiles reveal T-cell development, differentiation, and lineage-specific transcription factors including ZBTB25 as a novel NF-AT repressor. Blood 115, 5376-5384.
36. Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B., and Ideker, T. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome research 13, 2498-2504.
37. Oti, M., Snel, B., Huynen, M.A., and Brunner, H.G. (2006). Predicting disease genes using protein-protein interactions. J Med Genet 43, 691-698.
38. Sawcer, S. (2008). The complex genetics of multiple sclerosis: pitfalls and prospects. Brain 131, 3118-3131.
39. Tabor, H.K., Risch, N.J., and Myers, R.M. (2002). Candidate-gene approaches for studying complex genetic traits: practical considerations. Nature reviews Genetics 3, 391-397.
40. Hauser, M.A., Li, Y.J., Takeuchi, S., Walters, R., Noureddine, M., Maready, M., Darden, T., Hulette, C., Martin, E., Hauser, E., et al. (2003). Genomic convergence: identifying candidate genes for Parkinson's disease by combining serial analysis of gene expression and genetic linkage. Hum Mol Genet 12, 671-677.
41. Ban, M., McCauley, J.L., Zuvich, R., Baker, A., Bergamaschi, L., Cox, M., Kemppinen, A., D'Alfonso, S., Guerini, F.R., Lechner-Scott, J., et al. (2010). A non-synonymous SNP within membrane metalloendopeptidase-like 1 (MMEL1) is associated with multiple sclerosis. Genes and immunity 11, 660-664.

Figure Legends

Figure 1. Double Manhattan plot.

A Manhattan plot showing the gene-level p-values of both GWAS used in this study. Gene-level p-values from WTCCC2 are displayed at the top and those corresponding to meta2.5 at the bottom. A region in Chromosome 1 is enlarged to show detailed block structure. Blocks were defined as groups of contiguous genes with p-value of ≤ 0.05 (grayed area). The individual p-value of each gene is displayed as a colored circle ranging from green (not significant) to yellow to red (most significant). Note that the two plots are largely specular, denoting overall replication (see text).

Figure 2. Connectedness of first-order interaction networks.

The number of connections among significant genes were evaluated in the background of 1,000 random simulations (See text). **A.** The total number of edges was plotted as a function of the number of significant genes for each study. **B.** The size of the largest connected component is plotted as a function of the total number of edges. The colored lines represent 50^{th} (green), 90^{th} (yellow), 95^{th} (orange) and 99^{th} (red) percentiles obtained through simulations with random set of genes of similar size.

Figure 3. Intersection network.

Of the 118 nodes obtained by intersecting the resulting networks from each independent study, 88 were arranged in 13 sub-networks (ranging in size from 2 to 27) and 30 nodes remained isolated. Each node represents a gene product and each edge represents an experimental physical interaction reported in at least two independent publications. Thus, an edge is only displayed if the same interactions were identified in both studies. Isolated nodes in this representation may still have interactions within each of the studies, but they have not been preserved in both. White nodes are not significant. A color scale (yellow to red) denotes the significance of each node in WTCCC2. V-shaped nodes have nominal significant p-values in both studies. Nodes With a yellow outline denote genes containing bona-fide MS susceptibility variants. A different background color was chosen to highlight each of the 6 sub-networks with size ≥ 3 (sub-networks of size = 2 were grouped under the same background).

Figure 4. Proportion of validated discoveries using a network vs. a non-network approach

Of the 118 genes in the intersection network 88 genes were arranged in 13 sub-networks sizes 2-27. Of those, 54 genes were nominally significant in both studies. 55% of these genes were either bona-fide MS genes (24%) or fell into bona-fide MS blocks (31%). Of the 30 singletons from the 118 gene intersection network, 26 had significant p-values in both studies. 46% of these were either bona-fide MS genes (11%) or fell into bona-fide MS blocks (35%). From the 154 genes with significant p-values but not found in networks, only 25% were either bona-fide MS genes (8%) or fell into bona-fide MS blocks (17%).

Table 1. Gene-level significance, power, and network characteristics of each GWAS

Disease	Number of nominally significant genes	Area under ROC	Size of 1st order interaction net (nodes-edges)	Number of sub-networks (min-max size)	Size of union net (nodes-edges)	Sensitivity (%)	Specificity (%)
MS (WTCCC2)	1,996	0.95	838-401	15 (6-200)	464-820	88	56
MS (meta2.5)	1,706	0.88	761-304	16 (5-189)	605-1031	-	-
T1D	1,056	0.65	474-161	13 (5-170)	378-669	30	4.4
T2D	913	0.71	405-74	8 (10-211)	332-562	34	17.0
RA	937	0.66	360-68	16 (6-207)	347-632	12	1.7
CD	997	0.72	469-116	15 (5-231)	449-1066	88	16.0
CAD	831	0.60	393-75	15 (6-183)	299-491	52	8.3
HT	813	0.64	349-40	13 (6-167)	355-500	38	2.6

MS: multiple sclerosis [MIM 126200]; T1D: type 1 diabetes [MIM 222100]; T2D: type 2 diabetes [MIM 125853]; RA: rheumatoid arthritis [MIM 180300]; CD: Crohn's disease [MIM 266600]; CAD: coronary artery disease [MIM 607339]; HT: Hypertension [MIM 145500].

Table 2. Nominally significant genes in WTCCC2 and meta2.5 and arranged in networks in at least one study

Bona-fide MS susceptibility loci						Previously unreported (candidate) loci		
Gene			Block					
Gene symbol	p-value (WTCCC2)	p-value (meta2.5)	Gene symbol	p-value (WTCCC2)	p-value (meta2.5)	Gene symbol	p-value (WTCCC2)	p-value (meta2.5)
CD58	2E-05	1E-06	*TNFRSF14*	1E-07	0.00055	*PHGDH*	0.0008	0.00012
MERTK	0.0026	0.01698	*S1PR1*[a]	0.0261	0.00992	*ETS1*	0.0122	0.02008
IL12A	3E-05	0.00066	*GOLGB1*	1E-07	0.00197	*TRAF3*	0.0013	0.00113
IL7R	0.0005	0.00137	*KIF5A*	0.0002	0.00337	*BCL10*	2E-05	0.00054
IL12B	4E-06	1E-07	*CIITA*	1E-07	0.0129	*CD48*	9E-06	0.00902
IL7	3E-05	0.00339	*SOCS1*	1E-07	2E-06	*REL*	0.0003	0.00047
IL2RA	0.0015	0.00063	*RBM17*	1E-05	0.00271	*C17orf57*	0.0117	0.00259
TNFRSF1A	1E-07	0.00019	*SCNN1A*	2E-06	0.00017	*KPNB1*	0.0002	1E-07
STAT3	0.0001	0.00004	*LTBR*	2E-06	0.00061	*CHERP*	0.0018	1.6E-05
MALT1	0.0002	0.00068	*CD27*	0.0003	0.0125	*TEC*	0.0007	0.00048
CD40	0.0002	0.02198	*VAMP1*	0.0003	0.01662	*CSF2*	0.0112	0.00755
MAPK1	5E-05	0.00063	*STAT5A*	0.0002	1.8E-05	*IRF1*	0.0036	0.01076
SCO2	1E-05	0.00315	*STAT5B*	0.0349	0.00099	*EIF3B*	5E-05	0.03787
VCAM1	**0.0008**	**0.00917**	*CLTC*	0.0066	0.00088	*JAK2*	0.01	0.01543

RGS1		**5E-06**	**0.00076**	*KEAP1*[b]		0.0019	0.03798	*PAX5*	0.0232	0.04573
TNFSF14		**0.0002**	**0.00452**	*PFDN4*		0.0001	0.01813	*RIC8A*	0.0061	0.02786
-	-		-	*TOP3B*		0.0006	0.00695	*NR1H3*	0.0007	0.00472
-	-		-	***BCL2L11***		**0.0005**	**0.00101**	*SART1*	0.0001	0.00115
-	-		-	***HCLS1***		**0.0035**	**0.0157**	*VPS33A*	0.0091	0.00124
-	-		-	***CDK4***		**1E-06**	**1.1E-05**	*MARK3*	0.0066	9.6E-05
-	-		-	***PITPNM2***		**0.0002**	**0.00133**	*FBF1*	0.0232	0.00657
-	-		-	***OGFOD2***		**0.0006**	**0.0032**	*PIK3R2*	1E-07	0.01093
-	-		-	***C3***		**0.0003**	**0.02719**	*PITPNB*	0.0059	0.01738
-	-		-	***PDE4A***		**1E-07**	**0.0048**	***BBC3***	**0.0006**	**0.01007**
-	-		-	***RTEL1***		**4E-05**	**0.00665**	***SLAMF1***	**0.008**	**0.0497**
-	-		-	***PPM1F***		**0.0006**	**0.00458**	***MORF4L1***	**0.0105**	**0.0022**
-	-		-	-	-		-	***MED26***	**0.001**	**0.0002**
-	-		-	-	-		-	***TXK***	**0.0002**	**0.00039**
-	-		-	-	-		-	***IKZF1***	**0.0068**	**0.04215**
-	-		-	-	-		-	***BMI1***	**0.01**	**0.0102**
-	-		-	-	-		-	***PSMD13***	**0.0054**	**0.02084**
-	-		-	-	-		-	***FOSL1***	**0.007**	**0.02244**
-	-		-	-	-		-	***CLIP1***	**0.005**	**4.6E-05**

-	-	-	-	-	-	***ITGAX***	**0.0025**	**0.01925**
-	-	-	-	-	-	***JUP***	**0.0306**	**0.01387**
-	-	-	-	-	-	***DNAJC7***	**0.0069**	**0.01956**
-	-	-	-	-	-	***RAB3A***	**1E-05**	**0.01381**

[a] Is significant but gene falls in a contiguous block. [b] Is significant but block in meta2.5 is smaller than in WTCCC2. Genes in bold are those with significant p-values in both studies but arranged as singletons in the intersection network from Figure 3

Figure 1
Click here to download high resolution image

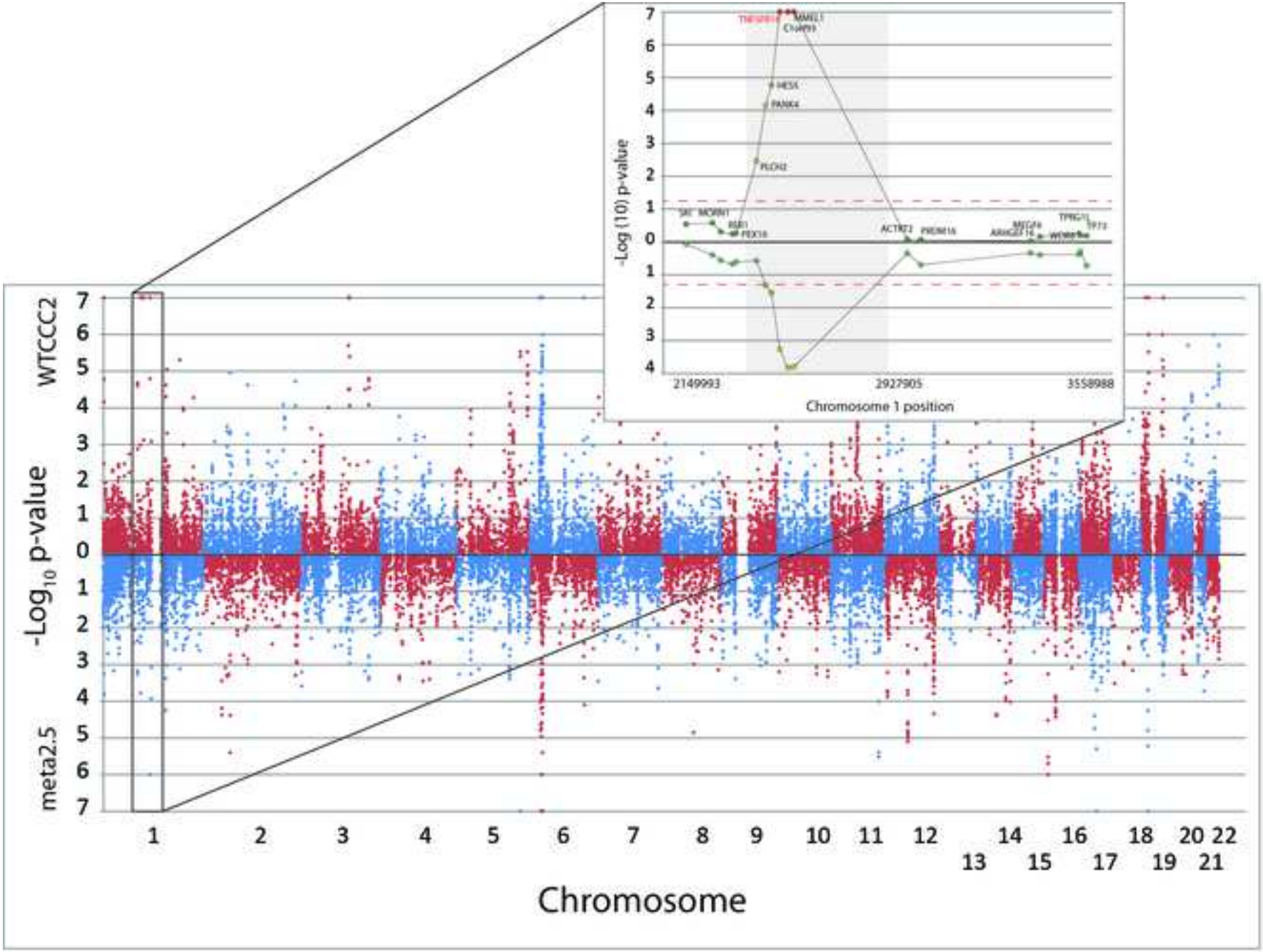


Figure 1

Figure 2
Click here to download high resolution image

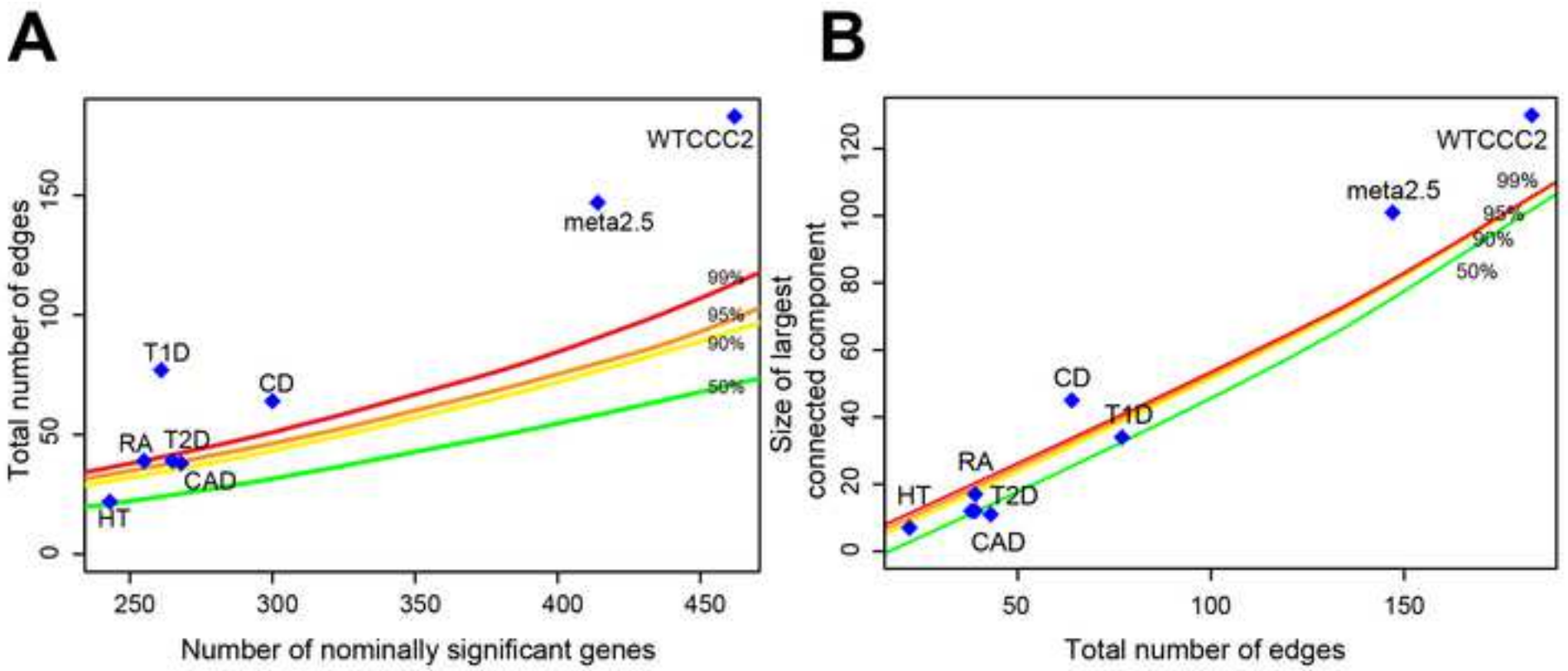


Figure 2

Figure 3
Click here to download high resolution image

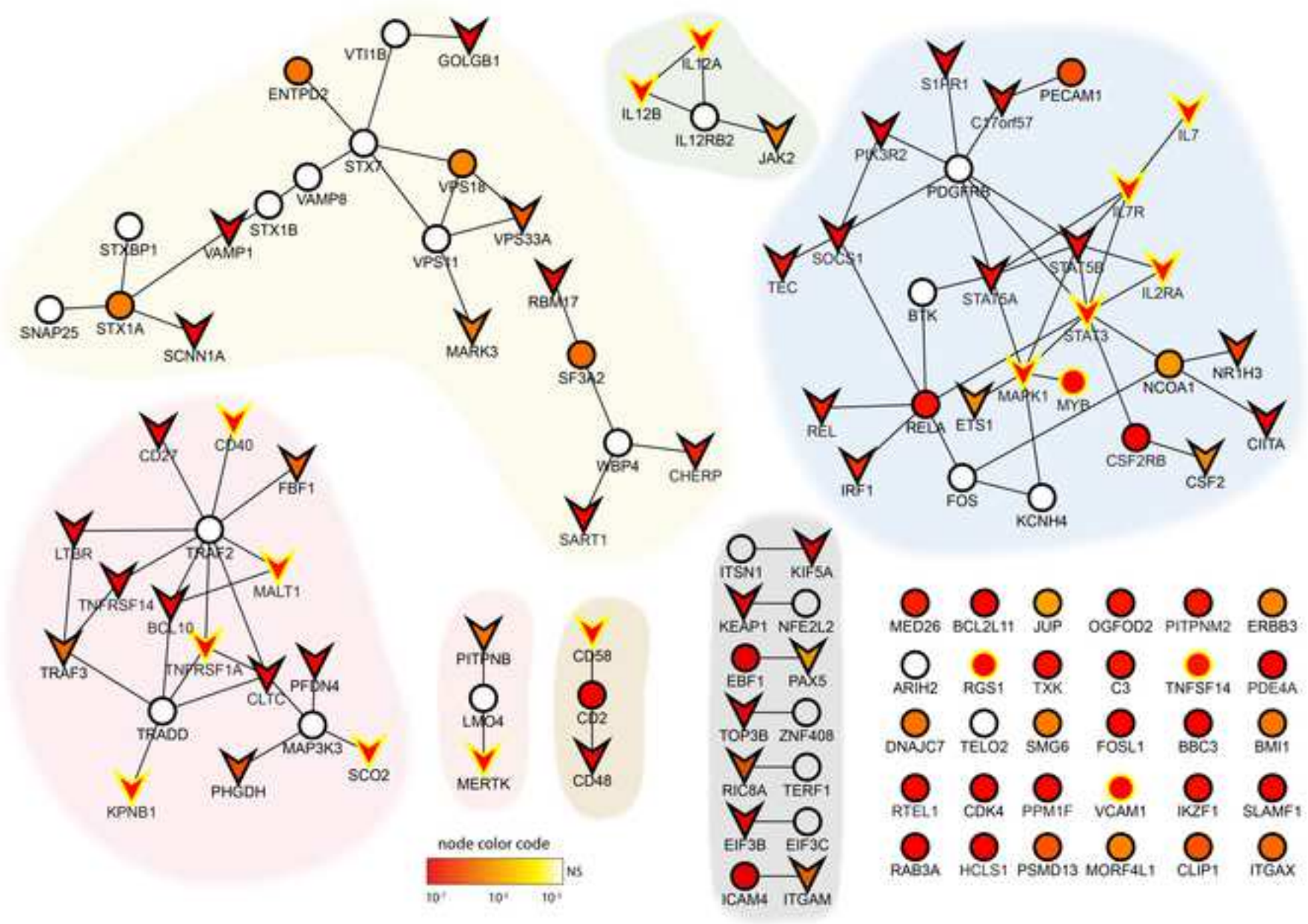


Figure 3

Figure 4
Click here to download high resolution image

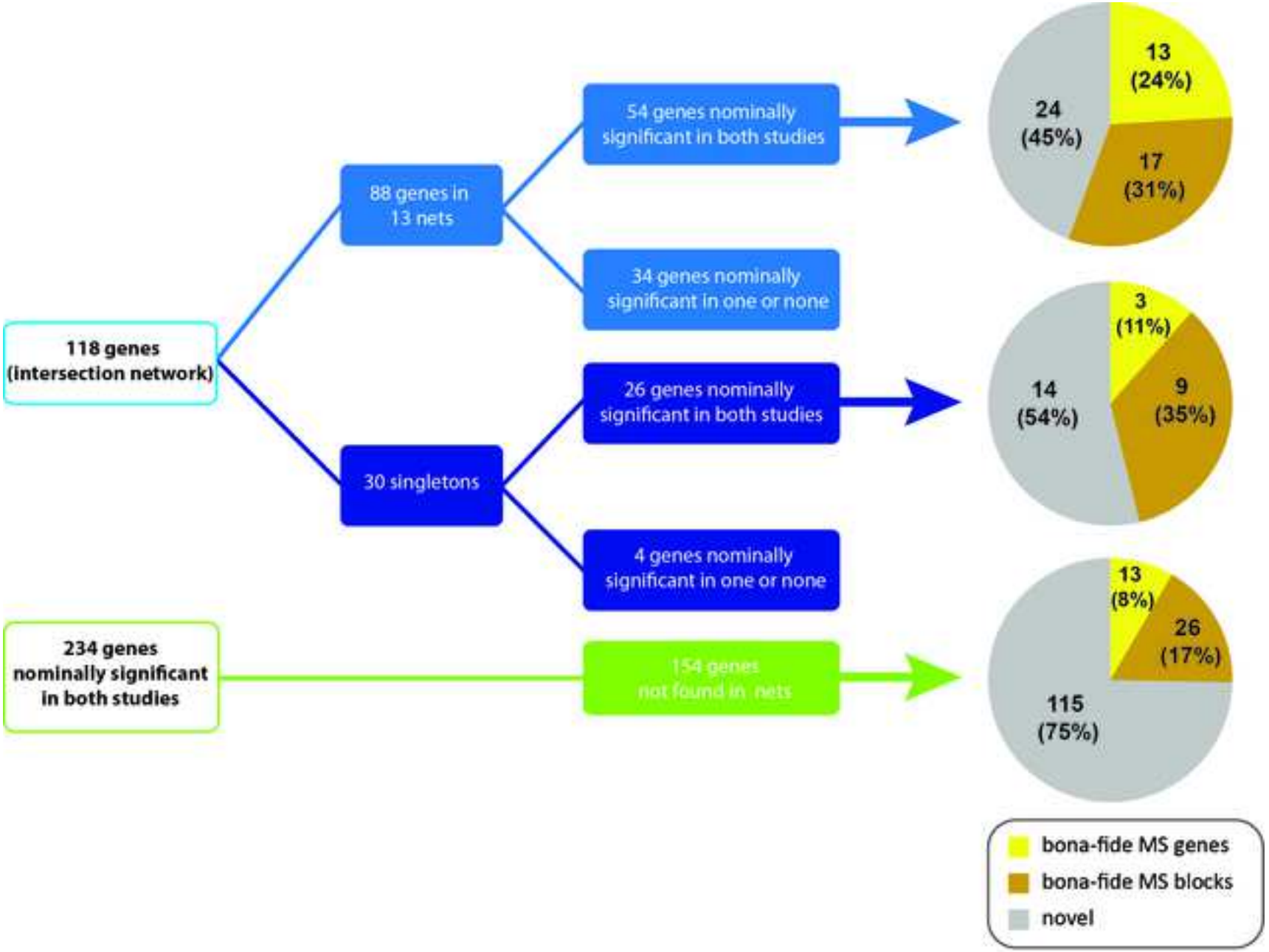


Figure 4

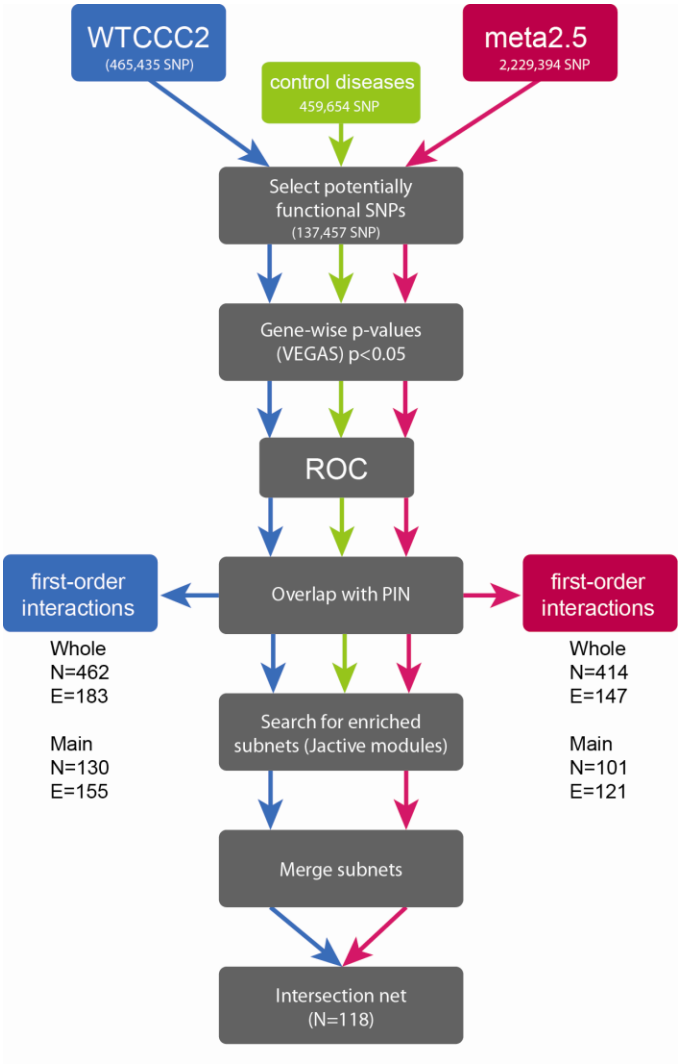
Network-based pathway analysis in multiple sclerosis with GWAS data from 15,000 cases and 30,000 controls

The International Multiple Sclerosis Genetics Consortium (IMSGC).

Supplemental Material

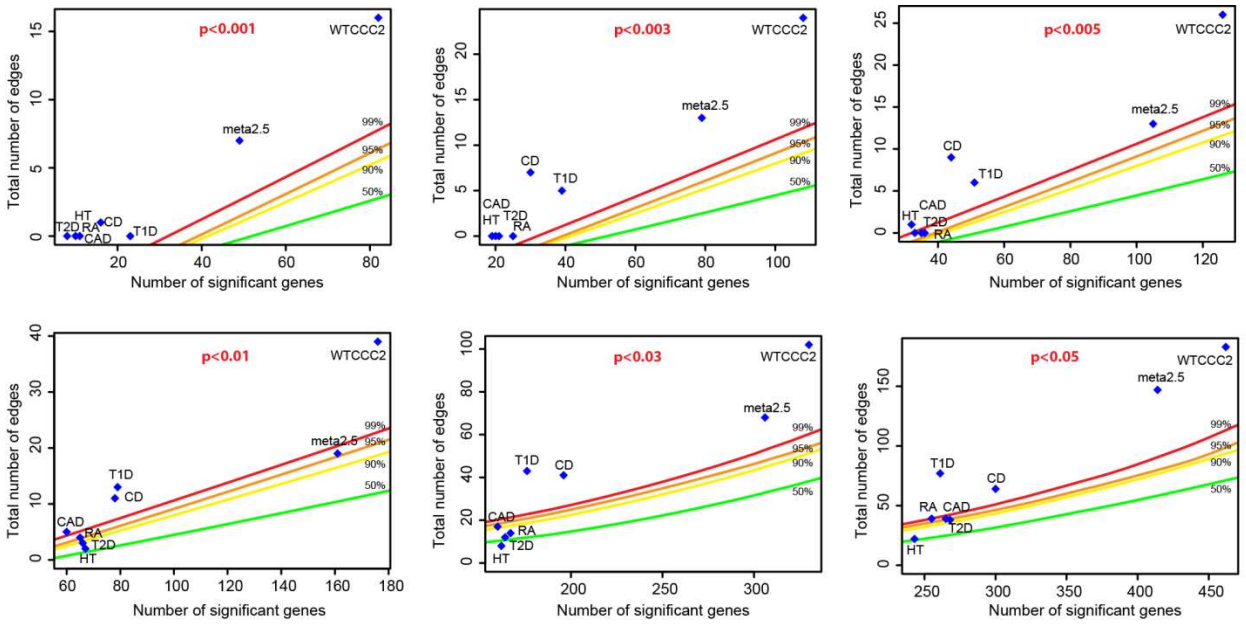
Contents

Supplementary Figure 1. Strategy



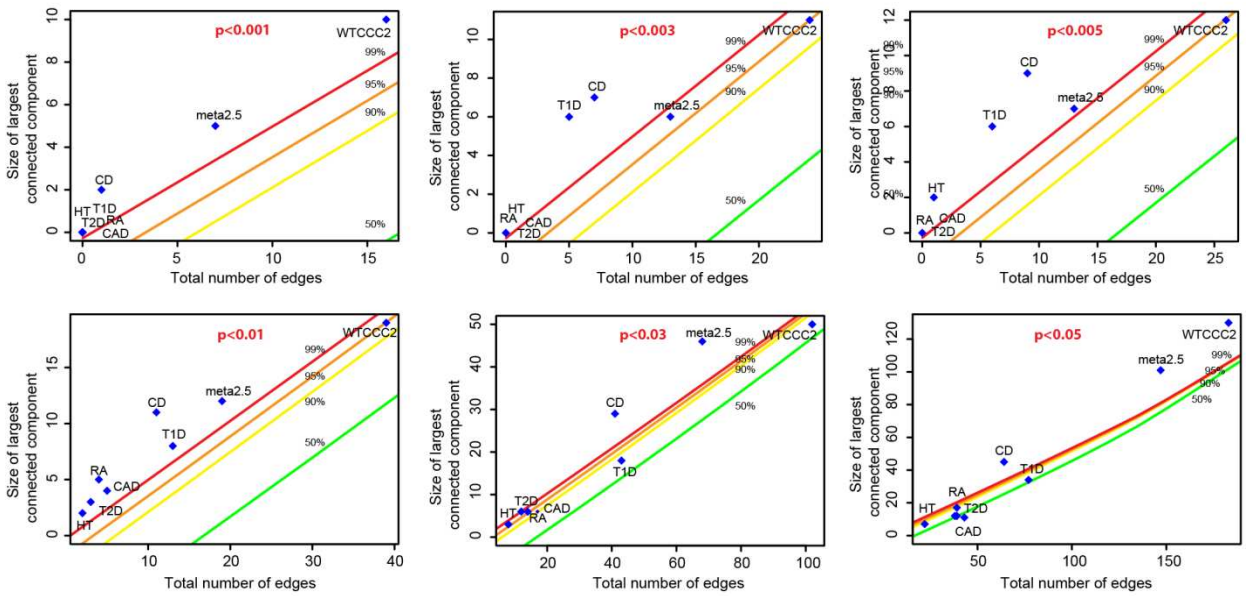
Summary-level data from two independent studies in MS were used, along with that of 6 diseases from the WTCCC1 study as controls. In the first step, SNPs with potentially functional consequences were selected. These included coding non-synonymous substitutions, and those mapping to 3' or 5' UTRs, and transcription factor or histone binding sites. Next, a gene-wise p-value was obtained using VEGAS, and genes with a p-value of 0.05 were analyzed further. ROC curves were computed to evaluate the power of each study to identify all known susceptibility loci (bona-fide) to date. In the following step gene p-values were overlapped with a curated protein interaction network (See methods) and interactions among nominally significant genes (first-order) were computed. With the goal of discovering additional associations, a heuristic search for sub-networks enriched in significant genes was conducted using jActive modules within Cytoscape. All significant sub-networks within each independent study were merged and finally the intersection network was computed, resulting in 118 nodes which were analyzed in depth.

Supplementary Figure 2. Total number of edges as a function of the number of significant number of genes at different significance thresholds



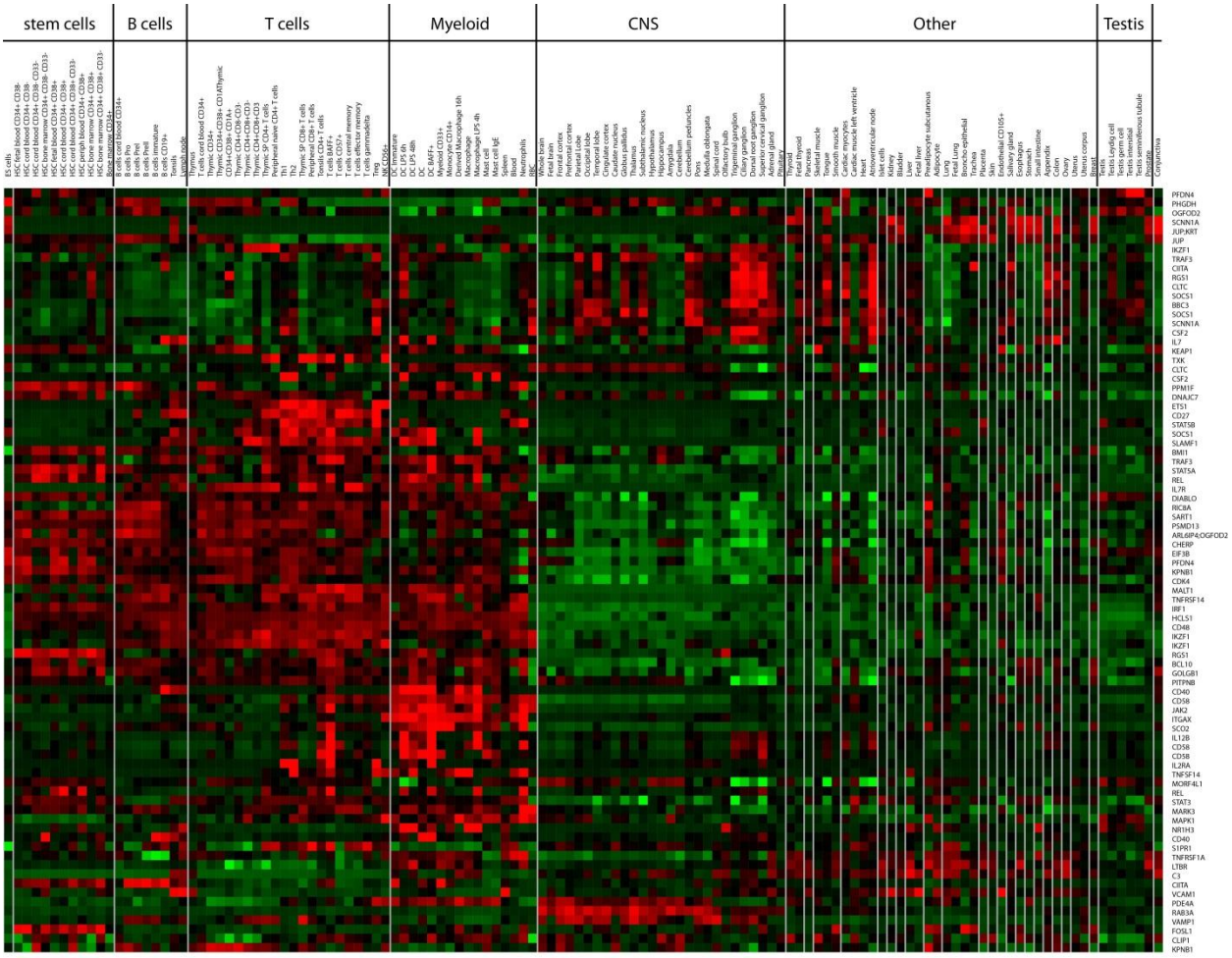
The total number of edges as a function of the number of significant number of genes at different significance thresholds is shown. At higher thresholds of significance (0.001) all diseases show more edges than expected by chance. As the significance drops, fewer studies (mostly those well powered) show this property. Notably, networks from WTCCC2 and meta2.5 are clearly more connected than chance

Supplementary Figure 3. Size of the largest connected component as a function of the total number of edges in the first order networks at different thresholds of significance



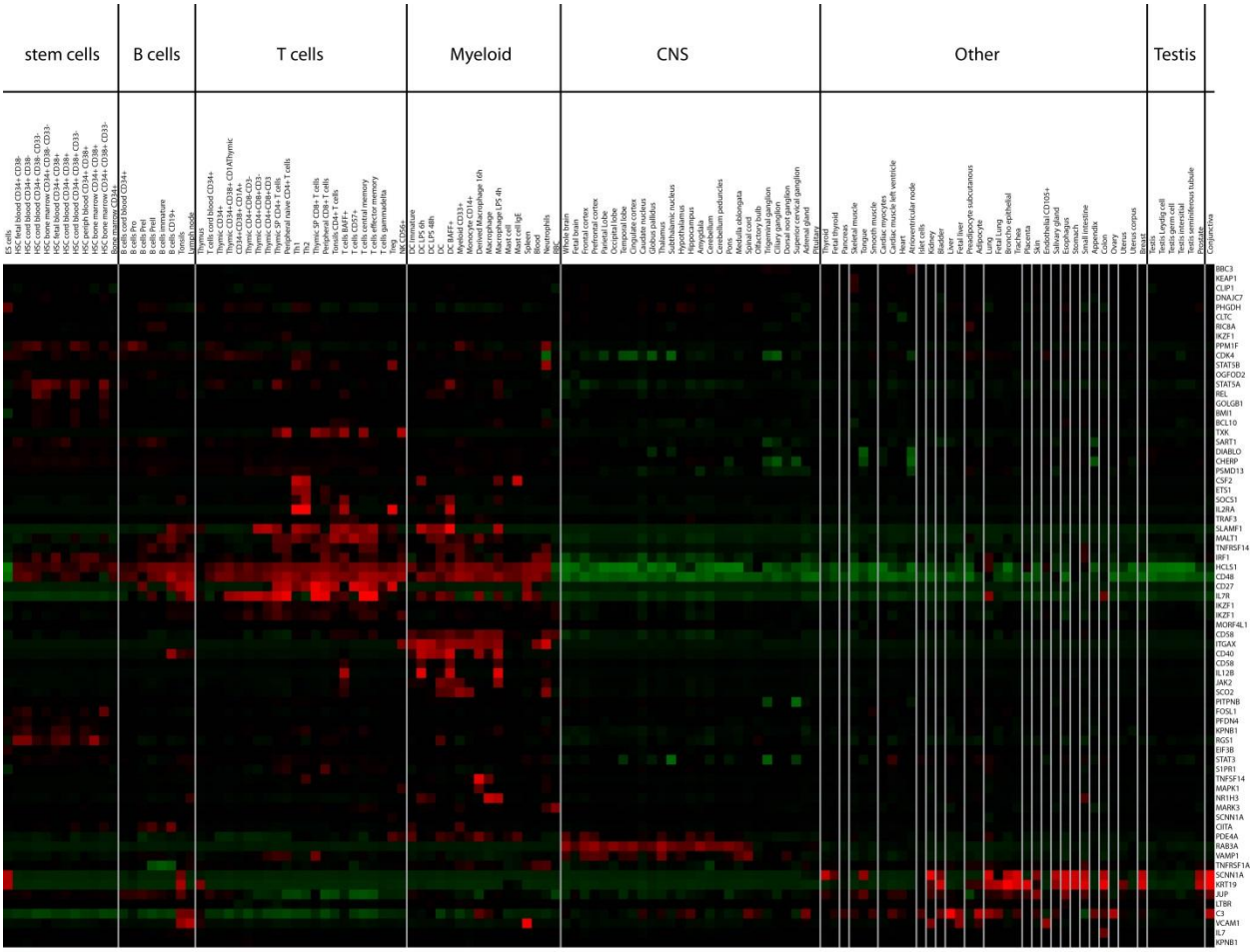
The size of the largest connected component as a function of the total number of edges in the first order networks at different thresholds of significance is shown. This Figure highlights that a large proportion of the total number of edges counts towards formation of the largest connected component of the first order interaction network. As expected, this is more evident at higher thresholds, but it is also true at 0.05 for WTCCC2 and meta2.5

Supplementary Figure 4. Transcript expression heatmap.



Transcript expression with the 88 genes from the intersection network which were arranged in networks was computed using the Gene enrichment profiler tool 35. This approach computes the expression of all mapping transcripts in a user-initiated query in each of 126 normal tissues. Almost two-thirds of genes in this list are highly expressed in immune-related tissues and cell types. Almost half are also highly expressed in the CNS and other tissues. Colors correspond to the level of enrichment in each tissue (green = depleted, black = no enrichment, red = enriched).

Supplementary Figure 5. Transcript enrichment heatmap



Transcript enrichment with the 88 genes from the intersection network which were arranged in networks was computed using the Gene enrichment profiler tool [35]. This approach computes the enrichment of all mapping transcripts in a user-initiated query in each of 126 normal tissues. A significant proportion of genes in this list, are enriched in immune-related tissues and cell types (specifically B, T and myeloid cells). Three genes (PDE4EA, RAB3A and VAMP1) appear to be enriched in CNS tissues. Colors correspond to the level of enrichment in each tissue (green = depleted, black = no enrichment, red = enriched).

Supplementary Table 1.

Supplementary Table 1. Study details

Disease	Number of cases	Number of controls	Reference
MS (WTCCC2)	9,772	17,376	{Sawcer, 2011 #7199}
MS (meta2.5)	5,545	12,153	{Patsopoulos, 2011 #6087}
T1D	2000	3,0000	{WTCCC, 2007 #1670}
T2D	2000	3,0000	{WTCCC, 2007 #1670}
RA	2000	3,0000	{WTCCC, 2007 #1670}
CD	2000	3,0000	{WTCCC, 2007 #1670}
CAD	2000	3,0000	{WTCCC, 2007 #1670}
HT	2000	3,0000	{WTCCC, 2007 #1670}

Supplementary Table 2. Prioritization of candidate genes within a block (Functional mapping)

			WTCCC2							Meta2.5						
Gene	Chr	position	nSNPs	Test	Pvalue	Best.SNP	SNP.pvalue	block#	block best pvalue	nSNPs	Test	Pvalue	Best.SNP	SNP.pvalue	block #	block best pvalue
PLCH2	1	2397613	9	35.00558	0.0034	rs17373634	0.000274	1	0.0000001	NA	NA	NA	NA	NA	NA	NA
PANK4	1	2429834	10	65.5516	6.90E-05	rs6667605	5.16E-07	1	0.0000001	10	23.06865	0.04864	rs6667605	0.000413	4	0.000149
HES5	1	2450043	7	60.87558	1.70E-05	rs6667605	5.16E-07	1	0.0000001	7	20.72394	0.02821	rs6667605	0.000413	4	0.000149
TNFRSF14	1	2479150	8	153.8989	1.00E-07	rs3748816	2.25E-13	1	0.0000001	8	51.5609	0.000551	rs3748816	2.81E-05	4	0.000149
C1orf93	1	2508108	6	202.3974	1.00E-07	rs3890745	5.79E-14	1	0.0000001	6	59.12327	0.000149	rs3748816	2.81E-05	4	0.000149
MMEL1	**1**	**2511940**	**6**	**202.3974**	**1.00E-07**	**rs3890745**	**5.79E-14**	**1**	**0.0000001**	**6**	**59.12327**	**0.000158**	**rs3748816**	**2.81E-05**	**4**	**0.000149**
VCAM1	**1**	**100957884**	**10**	**40.48064**	**0.000756**	**rs11586136**	**8.34E-06**	**32**	**0.0000001**	**10**	**28.67992**	**0.00917**	**rs2050471**	**0.002007**	**32**	**0.00917**
EXTL2	1	101110528	7	86.51482	1.00E-07	rs11809572	3.97E-09	32	0.0000001	7	22.15338	0.01564	rs12048904	0.000577	32	0.00917
SLC30A7	1	101134265	13	225.9379	1.00E-07	rs11809572	3.97E-09	32	0.0000001	13	39.0039	0.02525	rs12048904	0.000577	32	0.00917
DPH5	1	101227768	8	172.2086	1.00E-07	rs7521424	4.12E-09	32	0.0000001	8	20.38018	0.07685	rs7521424	0.03806	NA	0.07685
S1PR1*	1	101474892	16	38.97881	0.02608	rs2031959	0.01244	32	0.0000001	16	47.31656	0.00992	rs4459133	0.001194	33	0.00992
CD58	**1**	**116858679**	**15**	**97.04105**	**1.60E-05**	**rs1335532**	**2.01E-09**	**35**	**0.0000001**	**15**	**113.3989**	**1.00E-06**	**rs1335532**	**1.51E-09**	**36**	**0.000001**
IGSF3	1	116918553	6	68.51176	1.00E-07	rs1335532	2.01E-09	35	0.0000001	6	71.22613	1.00E-06	rs1335532	1.51E-09	36	0.000001
CD2	1	117098608	6	14.3167	0.04127	rs798037	0.01902	35	0.0000001	NA	NA	NA	NA	NA	NA	NA
ACOXL	2	111206679	13	55.90904	0.0016	rs1439287	9.41E-05	93	0.000024	14	41.65543	0.01134	rs1837369	0.0007737	96	0.001007
BCL2L11	2	111594961	18	122.7645	0.000528	rs1439287	9.41E-05	93	0.000024	18	110.5492	0.001007	rs9308742	0.0003182	96	0.001007
ANAPC1	2	112243110	7	56.36742	2.40E-05	rs17174870	7.78E-06	93	0.000024	7	36.1738	0.001386	rs17174870	0.000391	96	0.001007
MERTK	**2**	**112372661**	**23**	**116.6903**	**0.0026**	**rs17174870**	**7.78E-06**	**93**	**0.000024**	**23**	**75.72197**	**0.01698**	**rs11884641**	**0.0002625**	**96**	**0.001007**
TMEM87B	2	112529284	14	57.90995	0.00158	rs11884641	1.37E-05	93	0.000024	NA	NA	NA	NA	NA	NA	NA
SCHIP1	3	160474237	NA	NA	NA	NA	NA	NA	NA	34	85.7023	0.01683	rs7635938	0.004963	137	0.000657
IL12A	**3**	**161189322**	**14**	**102.9125**	**3.20E-05**	**rs2243123**	**3.68E-06**	**159**	**0.000032**	**14**	**73.1439**	**0.000657**	**rs7610160**	**2.07E-05**	**137**	**0.000657**
FBXO40	3	122794859	13	30.59088	0.03828	rs7153	0.000212	153	0.0000001	13	28.83312	0.04871	rs7153	0.001342	129	0.001069
HCLS1	3	122832935	11	43.51908	0.00347	rs3732410	0.000172	153	0.0000001	11	33.69831	0.0157	rs7153	0.001342	129	0.001069
GOLGB1	3	122864737	13	148.2947	1.00E-07	rs6808500	9.61E-08	153	0.0000001	14	66.21549	0.001969	rs7153	0.001342	129	0.001069

IQCB1	3	122971299	10	206.7136	1.00E-07	rs1920298	6.35E-08	153	0.0000001	11	68.01742	0.002043	rs1920298	0.000972	129	0.001069
EAF2	3	123036723	10	170.883	1.00E-07	rs1920298	6.35E-08	153	0.0000001	11	57.20001	0.00271	rs1920298	0.000972	129	0.001069
SLC15A2	3	123095976	9	91.98277	3.10E-05	rs4285028	3.70E-08	153	0.0000001	9	49.99358	0.00277	rs9289185	0.000415	129	0.001069
ILDR1	3	123188859	12	91.78509	4.00E-06	rs4285028	3.70E-08	153	0.0000001	12	52.89139	0.001069	rs9289185	0.000415	129	0.001069
CD86	**3**	**123256910**	**11**	**70.89507**	**8.70E-05**	**rs9282641**	**1.54E-09**	**153**	**0.0000001**	**11**	**41.39866**	**0.00473**	**rs9282641**	**2.77E-05**	**129**	**0.001069**
CASR	3	123385219	NA	NA	NA	NA	NA	NA	NA	16	41.75711	0.01922	rs1979869	0.002792	129	0.001069
SPEF2	5	35653745	14	80.14992	0.00019	rs11748781	0.00066	197	0.000118	15	59.49978	0.002249	rs931555	0.002935	181	0.000555
IL7R	**5**	**35892747**	**15**	**95.10324**	**0.000525**	**rs6897932**	**2.63E-06**	**197**	**0.000118**	**15**	**83.30192**	**0.00137**	**rs6897932**	**0.0002259**	**181**	**0.000555**
CAPSL	5	35940154	26	161.6188	0.000118	rs6897932	2.63E-06	197	0.000118	26	135.3675	0.000555	rs6897932	0.0002259	181	0.000555
UGT3A1	5	35988966	16	81.84978	0.001044	rs1993879	9.52E-06	197	0.000118	16	64.1848	0.00474	rs860413	0.0003444	181	0.000555
EBF1	5	158055500	26	69.06629	0.01174	rs10515789	0.00032	218	0.000003	NA	NA	NA	NA	NA	NA	NA
RNF145	5	158516996	10	41.9753	0.0038	rs6556396	0.00947	218	0.000003	NA	NA	NA	NA	NA	NA	NA
UBLCP1	5	158622666	7	65.14022	3.00E-06	rs2546890	2.74E-07	218	0.000003	8	79.78664	1E-07	rs2546890	7.95E-08	203	0.0000001
IL12B	**5**	**158674368**	**6**	**56.23183**	**4.00E-06**	**rs2546890**	**2.74E-07**	**218**	**0.000003**	**7**	**82.49851**	**1E-07**	**rs2546890**	**7.95E-08**	**203**	**0.0000001**
ALDH8A1	6	135280220	NA	NA	NA	NA	NA	NA	NA	10	47.16629	0.00167	rs9385708	8.06E-05	241	0.00167
HBS1L	6	135323213	NA	NA	NA	NA	NA	NA	NA	16	58.98443	0.01571	rs12526072	0.009443	241	0.00167
MYB #	**6**	**135544145**	**12**	**77.80818**	**0.00077**	**rs9321490**	**5.81E-06**	**258**	**0.0000001**	**12**	**8.631174**	**0.536**	**rs2050019**	**0.04948**	**NA**	**0.536**
AHI1	6	135646816	9	195.2345	1.00E-07	rs11154801	1.53E-12	258	0.0000001	10	74.40884	0.000435	rs11154801	0.0001161	242	0.000435
PKIA	8	79590890	10	40.70157	0.0141	rs11781853	0.00133	309	0.000033	10	117.5594	1.40E-05	rs10504676	1.38E-05	292	0.000014
FAM164A	8	79740884	6	92.28494	3.60E-05	rs3736169	3.44E-05	309	0.000033	6	58.15676	0.001025	rs1384804	0.0004862	292	0.000014
IL7	**8**	**79807559**	**4**	**44.15623**	**3.30E-05**	**rs1441438**	**5.09E-05**	**309**	**0.000033**	**4**	**23.16013**	**0.00339**	**rs2369440**	**0.0007909**	**292**	**0.000014**
IL2RA	**10**	**6093511**	**21**	**88.14574**	**0.00151**	**rs942201**	**2.58E-07**	**352**	**0.000014**	**21**	**100.1787**	**0.000627**	**rs11256497**	**5.91E-05**	**339**	**0.000627**
RBM17	10	6171012	19	145.3416	1.40E-05	rs942201	2.58E-07	352	0.000014	19	76.69935	0.00271	rs11256497	5.91E-05	339	0.000627
PFKFB3	10	6284900	23	46.51204	0.0307	rs630204	0.00416	352	0.000014	23	46.20471	0.03232	rs11257741	0.00899	339	0.000627
CD9	12	6179815	NA	NA	NA	NA	NA	NA	NA	16	58.56654	0.00934	rs2268016	0.003049	397	0.000132
PLEKHG6	12	6289862	16	128.8267	1.00E-07	rs1800693	1.84E-10	419	0.0000001	18	87.19434	0.000132	rs1800693	1.41E-05	397	0.000132
TNFRSF1A	**12**	**6308183**	**16**	**129.4852**	**1.00E-07**	**rs1800693**	**1.84E-10**	**419**	**0.0000001**	**18**	**83.38418**	**0.000191**	**rs1800693**	**1.41E-05**	**397**	**0.000132**
SCNN1A	12	6326273	18	130.4958	2.00E-06	rs1800693	1.84E-10	419	0.0000001	20	85.58752	0.000167	rs1800693	1.41E-05	397	0.000132
LTBR	12	6363617	14	95.2566	2.00E-06	rs1860545	1.60E-09	419	0.0000001	15	56.72396	0.000613	rs1860545	2.02E-05	397	0.000132
CD27	12	6424311	18	100.4861	0.000306	rs2250246	0.001324	419	0.0000001	19	57.0711	0.0125	rs2286596	0.004386	397	0.000132
TAPBPL	12	6431437	20	104.0096	0.000319	rs2250246	0.001324	419	0.0000001	21	57.9826	0.01617	rs2286596	0.004386	397	0.000132

VAMP1	12	6441664	20	104.0096	0.000326	rs2250246	0.001324	419	0.0000001	21	57.9826	0.01662	rs2286596	0.004386	397	0.000132
MRPL51	12	6471576	22	110.7377	0.000489	rs2250246	0.001324	419	0.0000001	23	56.23403	0.03407	rs2286596	0.004386	397	0.000132
NCAPD2	12	6473558	22	91.31248	0.001926	rs2532497	0.001606	419	0.0000001	NA	NA	NA	NA	NA	NA	NA
GAPDH	12	6513917	12	38.71814	0.01675	rs1048402	0.004681	419	0.0000001	NA	NA	NA	NA	NA	NA	NA
IFFO	12	6518954	12	38.71814	0.0159	rs1048402	0.004681	419	0.0000001	NA	NA	NA	NA	NA	NA	NA
MBD6	12	56202925	5	20.63647	0.02781	rs11172254	1.76E-05	437	0.0000001	NA	NA	NA	NA	NA	NA	NA
DCTN2	12	56210360	4	20.4288	0.01547	rs11172254	1.76E-05	437	0.0000001	NA	NA	NA	NA	NA	NA	NA
KIF5A	12	56230113	10	66.86074	0.000154	rs10083154	3.45E-06	437	0.0000001	10	43.49829	0.00337	rs923828	0.0001923	409	0.000008
PIP4K2C	12	56271323	7	64.86693	1.40E-05	rs10083154	3.45E-06	437	0.0000001	7	42.49269	0.00058	rs923828	0.0001923	409	0.000008
DTX3	12	56284870	7	64.86693	1.40E-05	rs10083154	3.45E-06	437	0.0000001	7	42.49269	0.000658	rs923828	0.0001923	409	0.000008
GEFT	12	56290229	7	64.86693	1.50E-05	rs10083154	3.45E-06	437	0.0000001	7	42.49269	0.000633	rs923828	0.0001923	409	0.000008
SLC26A10	12	56299959	9	109.5714	2.00E-06	rs774890	1.93E-06	437	0.0000001	9	65.53958	0.000196	rs923828	0.0001923	409	0.000008
B4GALNT1	12	56305817	9	110.5821	3.00E-06	rs774890	1.93E-06	437	0.0000001	9	72.28768	0.000151	rs923828	0.0001923	409	0.000008
OS9	12	56374152	13	191.5272	1.00E-07	rs12368653	1.99E-07	437	0.0000001	12	124.6172	1.40E-05	rs2069502	1.53E-05	409	0.000008
CENTG1	12	56405260	12	164.6265	3.00E-06	rs12368653	1.99E-07	437	0.0000001	11	120.2294	8.00E-06	rs2069502	1.53E-05	409	0.000008
TSPAN31	12	56425050	10	144.7957	1.00E-06	rs12368653	1.99E-07	437	0.0000001	9	107.3298	2.30E-05	rs2069502	1.53E-05	409	0.000008
CDK4	12	56428269	10	144.7957	1.00E-06	rs12368653	1.99E-07	437	0.0000001	9	107.3298	1.10E-05	rs2069502	1.53E-05	409	0.000008
MARCH9	12	56435166	11	157.938	1.00E-07	rs12368653	1.99E-07	437	0.0000001	10	118.9335	1.40E-05	rs2069502	1.53E-05	409	0.000008
CYP27B1	**12**	**56442383**	**11**	**157.938**	**1.00E-06**	**rs12368653**	**1.99E-07**	**437**	**0.0000001**	**10**	**118.9335**	**1.40E-05**	**rs2069502**	**1.53E-05**	**409**	**0.000008**
METTL1	12	56448617	10	137.9894	3.00E-06	rs12368653	1.99E-07	437	0.0000001	10	118.9335	1.60E-05	rs2069502	1.53E-05	409	0.000008
FAM119B	12	56452649	10	137.9894	5.00E-06	rs12368653	1.99E-07	437	0.0000001	10	118.9335	1.00E-05	rs2069502	1.53E-05	409	0.000008
TSFM	12	56462802	9	138.4423	4.00E-06	rs12368653	1.99E-07	437	0.0000001	9	113.2646	1.40E-05	rs2069502	1.53E-05	409	0.000008
AVIL	12	56477709	9	121.1612	6.00E-06	rs703842	4.76E-06	437	0.0000001	9	109.2607	1.30E-05	rs2069502	1.53E-05	409	0.000008
CTDSP2	12	56499976	6	80.25052	7.00E-06	rs10877013	7.77E-06	437	0.0000001	6	70.66554	2.70E-05	rs2291617	1.56E-05	409	0.000008
ABCB9	12	121979491	10	39.66879	0.00325	rs7296418	0.0001872	452	0.000227	11	31.52114	0.01559	rs937564	0.001011	429	0.00133
OGFOD2	12	122025306	5	36.72044	0.000574	rs7296418	0.0001872	452	0.000227	6	29.8209	0.0032	rs937564	0.001011	429	0.00133
ARL6IP4	12	122030832	4	35.86565	0.000398	rs7296418	0.0001872	452	0.000227	5	28.83928	0.00204	rs937564	0.001011	429	0.00133
PITPNM2	12	122033979	11	103.7941	0.000227	rs949143	0.0001538	452	0.000227	12	82.85517	0.00133	rs1106240	0.00068	429	0.00133
MPHOSPH9	**12**	**122206898**	**9**	**85.02801**	**0.000316**	**rs949143**	**0.0001538**	**452**	**0.000227**	**9**	**64.62894**	**0.001723**	**rs1790098**	**0.000653**	**429**	**0.00133**
C12orf65	12	122283415	4	34.99708	0.000528	rs4460848	0.0002456	452	0.000227	4	17.49441	0.01664	rs4460848	0.008892	429	0.00133
CDK2AP1	12	122311492	5	30.84473	0.002039	rs7304782	0.0005952	452	0.000227	5	19.05534	0.01919	rs1060105	0.003909	429	0.00133

SBNO1	12	122346407	5	19.0769	0.02449	rs1060105	0.002443	452	0.000227	NA	NA	NA	NA	NA	NA	NA
SETD8	12	122434656	1	4.27311	0.03928	rs1662	0.03872	452	0.000227	NA	NA	NA	NA	NA	NA	NA
RILPL2	12	122465888	1	4.27311	0.03768	rs1662	0.03872	452	0.000227	NA	NA	NA	NA	NA	NA	NA
U1SNRNPBP	12	122508603	1	4.27311	0.03892	rs1662	0.03872	452	0.000227	NA	NA	NA	NA	NA	NA	NA
CIITA	16	10878555	16	161.3609	1.00E-07	rs1035089	5.27E-13	508	0.0000001	16	43.75733	0.0129	rs1035089	0.004249	486	0.000001
DEXI	16	10930248	16	237.0363	1.00E-07	rs1035089	5.27E-13	508	0.0000001	16	59.22137	0.0046	rs7403919	0.001804	486	0.000001
CLEC16A	**16**	**10945942**	**57**	**1271.805**	**1.00E-07**	**rs7200786**	**6.30E-14**	**508**	**0.0000001**	**57**	**442.8161**	**3.00E-05**	**rs7200786**	**5.39E-06**	**486**	**0.000001**
SOCS1	16	11255774	21	225.7989	1.00E-07	rs12596260	2.21E-06	508	0.0000001	21	173.8217	2.00E-06	rs243324	1.33E-06	486	0.000001
TNP2	16	11269214	21	267.3649	1.00E-07	rs12928822	4.52E-07	508	0.0000001	21	200.8308	1.00E-06	rs243324	1.33E-06	486	0.000001
PRM3	16	11274644	21	267.3649	1.00E-07	rs12928822	4.52E-07	508	0.0000001	21	200.8308	1.00E-06	rs243324	1.33E-06	486	0.000001
PRM2	16	11276993	20	255.2985	1.00E-07	rs12928822	4.52E-07	508	0.0000001	20	192.3831	3.00E-06	rs243324	1.33E-06	486	0.000001
PRM1	16	11282193	18	243.1464	1.00E-07	rs12928822	4.52E-07	508	0.0000001	18	177.7196	1.00E-06	rs243324	1.33E-06	486	0.000001
C16orf75	16	11346811	13	116.5316	1.50E-05	rs12928822	4.52E-07	508	0.0000001	13	78.72053	0.000553	rs12928822	8.66E-05	486	0.000001
LITAF	16	11549356	8	31.6959	0.002179	rs7189692	0.001555	508	0.0000001	NA	NA	NA	NA	NA	NA	NA
GHDC	17	37594631	NA	NA	NA	NA	NA	NA	NA	2	7.441067	0.03056	rs7214921	0.01134	527	0.000018
STAT5B	17	37604720	5	16.57455	0.03492	rs8074524	0.02485	545	0.000132	5	37.56669	0.000989	rs8074524	0.000246	527	0.000018
STAT5A	17	37693090	7	68.10787	0.000219	rs9891119	4.64E-07	545	0.000132	7	88.88289	1.80E-05	rs9891119	2.71E-06	527	0.000018
STAT3	**17**	**37718868**	**12**	**90.7947**	**0.000132**	**rs9891119**	**4.64E-07**	**545**	**0.000132**	**12**	**101.0319**	**4.00E-05**	**rs9891119**	**2.71E-06**	**527**	**0.000018**
PTRF	17	37808000	9	54.90192	0.000277	rs9891119	4.64E-07	545	0.000132	9	50.48688	0.000609	rs9891119	2.71E-06	527	0.000018
ATP6V0A1	17	37864387	9	50.70907	0.00135	rs8067384	0.0006386	545	0.000132	9	50.48688	0.000533	rs9891119	2.71E-06	527	0.000018
NAGLU	17	37941476	5	34.27517	0.00173	rs8067384	0.0006386	545	0.000132	NA	NA	NA	NA	NA	NA	NA
HSD17B1	17	37957509	4	23.07523	0.002021	rs8067384	0.0006386	545	0.000132	NA	NA	NA	NA	NA	NA	NA
COASY	17	37967617	5	17.14396	0.01955	rs8067384	0.0006386	545	0.000132	NA	NA	NA	NA	NA	NA	NA
MLX	17	37972603	6	17.54413	0.04177	rs8067384	0.0006386	545	0.000132	NA	NA	NA	NA	NA	NA	NA
YPEL2	17	54763834	NA	NA	NA	NA	NA	NA	NA	16	41.2151	0.03293	rs2290266	0.002976	533	0.000327
DHX40	17	54997667	NA	NA	NA	NA	NA	NA	NA	5	34.95887	0.002139	rs7210832	0.000292	533	0.000327
CLTC	17	55051831	9	40.14034	0.00658	rs7215180	5.61E-05	553	0.000001	9	56.41087	0.000945	rs907066	0.000243	533	0.000327
PTRH2	17	55129448	7	50.67724	0.000483	rs2777899	4.13E-06	553	0.000001	7	53.20011	0.000327	rs907066	0.000243	533	0.000327
TMEM49	17	55139644	11	93.71344	3.10E-05	rs2777899	4.13E-06	553	0.000001	11	68.45955	0.000436	rs907066	0.000243	533	0.000327
TUBD1	17	55291632	5	78.84267	2.00E-06	rs1292053	7.47E-06	553	0.000001	5	24.08787	0.01027	rs1292034	0.001932	533	0.000327
RPS6KB1	**17**	**55325224**	**8**	**112.4098**	**1.00E-06**	**rs180515**	**1.37E-07**	**553**	**0.000001**	**8**	**36.00879**	**0.00815**	**rs1292034**	**0.001932**	**533**	**0.000327**

RNFT1	17	55384504	6	76.59226	3.00E-06	rs180515	1.37E-07	553	0.000001	6	26.52816	0.00679	rs1292034	0.001932	533	0.000327
NEDD4L	18	53862777	64	152.0905	0.00801	rs9965182	0.0001464	548	0.000684	NA	NA	NA	NA	NA	NA	NA
ALPK2	18	54299461	62	115.8401	0.03872	rs11659233	6.55E-05	548	0.000684	59	119.8022	0.02169	rs12456021	7.81E-06	570	0.000207
MALT1	**18**	**54489597**	**24**	**95.42968**	**0.000684**	**rs4545915**	**0.0001299**	**548**	**0.000684**	**24**	**110.3439**	**0.000207**	**rs7238078**	**2.20E-06**	**570**	**0.000207**
MRPL4	19	10223639	15	41.75291	0.01599	rs2569693	0.001163	587	0.0000001	NA	NA	NA	NA	NA	NA	NA
ICAM1	19	10242516	19	70.74362	0.00131	rs2278442	0.0002943	587	0.0000001	NA	NA	NA	NA	NA	NA	NA
ICAM4	19	10258649	17	68.57281	0.000749	rs2278442	0.0002943	587	0.0000001	NA	NA	NA	NA	NA	NA	NA
ICAM5	19	10261654	19	82.45249	0.000275	rs2278442	0.0002943	587	0.0000001	NA	NA	NA	NA	NA	NA	NA
GLP-1	19	10276478	18	81.82661	0.000233	rs2278442	0.0002943	587	0.0000001	NA	NA	NA	NA	NA	NA	NA
FDX1L	19	10281890	17	81.24666	0.000213	rs2278442	0.0002943	587	0.0000001	NA	NA	NA	NA	NA	NA	NA
RAVER1	19	10287888	20	86.30058	0.000256	rs2278442	0.0002943	587	0.0000001	NA	NA	NA	NA	NA	NA	NA
ICAM3	19	10305451	19	85.59165	0.000201	rs2278442	0.0002943	587	0.0000001	NA	NA	NA	NA	NA	NA	NA
TYK2 %	**19**	**10322203**	**13**	**72.08147**	**3.70E-05**	**rs8112449**	**1.48E-06**	**587**	**0.0000001**	**NA**	**NA**	**NA**	**NA**	**NA**	**NA**	**NA**
CDC37	19	10362808	6	28.88318	0.002356	rs8112449	1.48E-06	587	0.0000001	NA	NA	NA	NA	NA	NA	NA
PDE4A	19	10392332	8	55.23604	1.00E-07	rs8112449	1.48E-06	587	0.0000001	8	26.02697	0.004797	rs7246953	0.003807	557	0.004797
KEAP1	19	10457795	8	32.06531	0.00185	rs7246953	4.59E-05	587	0.0000001	8	18.53504	0.03798	rs7246953	0.003807	557	0.004797
S1PR5	19	10484622	13	37.84331	0.01475	rs7246953	4.59E-05	587	0.0000001	13	39.38159	0.01214	rs7246953	0.003807	557	0.004797
GMEB2	20	61689398	10	31.73106	0.01743	rs6011002	0.0002361	640	0.000044	NA	NA	NA	NA	NA	NA	NA
STMN3	20	61741504	11	46.42023	0.00362	rs6011002	0.0002361	640	0.000044	NA	NA	NA	NA	NA	NA	NA
RTEL1	20	61760090	12	90.32903	4.40E-05	rs2427536	6.78E-05	640	0.000044	12	46.88628	0.00665	rs1151625	0.000537	591	0.00521
TNFRSF6B	**20**	**61796731**	**10**	**87.28124**	**4.80E-05**	**rs2427536**	**6.78E-05**	**640**	**0.000044**	**10**	**45.52191**	**0.00583**	**rs1151625**	**0.000537**	**591**	**0.00521**
ARFRP1	20	61801252	10	87.28124	5.30E-05	rs2427536	6.78E-05	640	0.000044	10	45.52191	0.00521	rs1151625	0.000537	591	0.00521
ZGPAT	20	61809237	11	87.77265	6.80E-05	rs2427536	6.78E-05	640	0.000044	11	45.82637	0.00692	rs1151625	0.000537	591	0.00521
LIME1	20	61838421	7	51.32375	0.00074	rs2427536	6.78E-05	640	0.000044	7	31.73595	0.0093	rs1151625	0.000537	591	0.00521
LIME1	20	61838421	7	51.32375	0.000714	rs2427536	6.78E-05	640	0.000044	7	31.90353	0.00877	rs1151625	0.000537	591	0.00521
SLC2A4RG	20	61841654	7	49.017	0.000889	rs2427536	6.78E-05	640	0.000044	8	33.6065	0.01039	rs1151625	0.000537	591	0.00521
ZBTB46	20	61846321	8	60.22284	0.000378	rs2427536	6.78E-05	640	0.000044	NA	NA	NA	NA	NA	NA	NA
MMP9	20	44070953	NA	NA	NA	NA	NA	NA	NA	5	14.44802	0.04832	rs3933239	0.03001	585	0.00534
SLC12A5	20	44091244	10	52.28552	0.000208	rs2425752	1.67E-06	630	0.000105	10	33.97052	0.00534	rs2425752	2.76E-05	585	0.00534
NCOA5	20	44123032	9	56.23764	0.000105	rs2425752	1.67E-06	630	0.000105	9	27.92941	0.01336	rs2425752	2.76E-05	585	0.00534
CD40	**20**	**44180312**	**8**	**52.65425**	**0.000179**	**rs2425752**	**1.67E-06**	**630**	**0.000105**	**8**	**24.02676**	**0.02198**	**rs2425752**	**2.76E-05**	**585**	**0.00534**

ZNF217	20	51617016	6	27.63551	0.001877	rs1555926	0.00088	636	0.000002	NA	NA	NA	NA	NA	NA	NA
BCAS1	20	51993485	42	94.39409	0.01318	rs10485442	0.00015	636	0.000002	NA	NA	NA	NA	NA	NA	NA
CYP24A1	**20**	**52203394**	**22**	**148.1899**	**2.00E-06**	**rs2248359**	**5.15E-09**	**636**	**0.000002**	**22**	**60.7498**	**0.00861**	**rs2248359**	**0.002694**	**588**	**0.00861**
PFDN4	20	52257908	13	87.4648	0.000149	rs2248359	5.15E-09	636	0.000002	13	40.99389	0.01813	rs2248359	0.002694	588	0.00861
PPIL2	22	20350272	20	51.05545	0.02358	rs8139619	0.001244	649	0.000045	NA	NA	NA	NA	NA	NA	NA
YPEL1	22	20381825	20	99.89337	0.000276	rs2283792	3.99E-06	649	0.000045	19	72.94745	0.00276	rs5999264	0.000551	605	0.000632
MAPK1	**22**	**20443946**	**16**	**125.0074**	**4.50E-05**	**rs2283792**	**3.99E-06**	**649**	**0.000045**	**15**	**91.61399**	**0.000632**	**rs5999264**	**0.000551**	**605**	**0.000632**
PPM1F	22	20603792	14	83.64046	0.000598	rs240066	4.56E-06	649	0.000045	15	61.75075	0.00458	rs240066	0.001107	605	0.000632
TOP3B	22	20641402	14	77.70595	0.000564	rs240066	4.56E-06	649	0.000045	15	54.12976	0.00695	rs240066	0.001107	605	0.000632
SBF1	22	49232100	14	41.6213	0.01286	rs131820	0.00024	665	0.000002	NA	NA	NA	NA	NA	NA	NA
ADM2	22	49266877	12	77.60083	1.20E-05	rs140522	3.85E-06	665	0.000002	12	38.464	0.00522	rs140522	0.004194	612	0.001406
MIOX	22	49272078	12	86.41554	7.00E-06	rs140522	3.85E-06	665	0.000002	12	46.99273	0.002274	rs140521	0.001953	612	0.001406
LMF2	22	49288245	11	88.94345	2.00E-06	rs140522	3.85E-06	665	0.000002	10	45.13558	0.001406	rs140521	0.001953	612	0.001406
NCAPH2	22	49293510	14	101.3098	1.80E-05	rs140522	3.85E-06	665	0.000002	13	52.71693	0.00301	rs140521	0.001953	612	0.001406
SCO2	**22**	**49308862**	**11**	**94.12118**	**1.20E-05**	**rs140522**	**3.85E-06**	**665**	**0.000002**	**10**	**46.72305**	**0.00315**	**rs140521**	**0.001953**	**612**	**0.001406**
TYMP	22	49311047	12	94.34466	1.60E-05	rs140522	3.85E-06	665	0.000002	11	46.72347	0.00363	rs140521	0.001953	612	0.001406
LOC440836	22	49315720	11	87.57769	2.40E-05	rs140522	3.85E-06	665	0.000002	10	45.38894	0.00396	rs140521	0.001953	612	0.001406
KLHDC7B	22	49333327	13	93.14997	8.00E-05	rs140522	3.85E-06	665	0.000002	12	47.40411	0.00739	rs140521	0.001953	612	0.001406
LOC644186	22	49336406	14	97.52122	8.70E-05	rs140522	3.85E-06	665	0.000002	13	49.7026	0.00807	rs140521	0.001953	612	0.001406
CHKB	22	49354155	13	71.70785	0.000501	rs140522	3.85E-06	665	0.000002	12	36.75617	0.01882	rs140521	0.001953	612	0.001406
CPT1B	22	49354155	13	71.70785	0.00049	rs140522	3.85E-06	665	0.000002	12	36.75617	0.0199	rs140521	0.001953	612	0.001406
MAPK8IP2	22	49385996	10	25.45299	0.03745	rs140518	0.00379	665	0.000002	NA	NA	NA	NA	NA	NA	NA

* S1PR1 is significant in both studies but falls in a contiguous block in meta2.5

MYB is borderline significant in meta2.5 (p=0.536)

% TYK2 is not significant in meta2.5

rows in red represent genes with p-values in both studies and identified in the intersection network (candidates)

Rows in bold represent bona-fide MS susceptibility genes

Supplementary Table 3. Candidate susceptibility variant-containing genes

			WTCCC2							Meta2.5						
Gene	Chr	Start	nSNPs	Test	Pvalue	Best.SNP	SNP.pvalue	block	block best pvalue	nSNPs	Test	Pvalue	Best.SNP	SNP.pvalue	block	block best pvalue
SYDE2	1	85395943	13	35.22858	0.02408	rs17382610	0.002596	28	0.000021	NA	NA	NA	NA	NA	NA	NA
C1orf52	1	85490243	20	135.824	2.50E-05	rs2030075	9.61E-06	28	0.000021	20	95.39334	0.000676	rs2030075	1.98E-05	27	0.000541
BCL10	1	85504047	21	137.7897	2.10E-05	rs2030075	9.61E-06	28	0.000021	21	99.62914	0.000541	rs2030075	1.98E-05	27	0.000541
DDAH1	1	85556755	33	139.9095	0.00133	rs2030075	9.61E-06	28	0.000021	33	110.6173	0.00451	rs2030075	1.98E-05	27	0.000541
ZNF697	1	1.2E+08	16	41.03661	0.01711	rs639216	0.00046	37	0.000823	15	62.25939	0.00083	rs639216	1.06E-05	38	0.000117
PHGDH	1	1.2E+08	19	82.24214	0.000823	rs639216	0.00046	37	0.000823	19	102.1858	0.000117	rs639216	1.06E-05	38	0.000117
HMGCS2	1	1.2E+08	18	54.77674	0.0092	rs592762	0.00528	37	0.000823	18	50.11705	0.01403	rs592762	1.75E-05	38	0.000117
SLAMF1	1	1.59E+08	13	38.75361	0.008	rs3795324	0.0004	44	0.000009	15	29.98571	0.0497	rs3795324	0.02557	49	0.00344
CD48	1	1.59E+08	12	62.79755	9.00E-06	rs983494	0.00015	44	0.000009	13	33.50673	0.00902	rs983494	0.001565	49	0.00344
SLAMF7	1	1.59E+08	8	38.8205	0.000499	rs983494	0.00015	44	0.000009	8	30.0873	0.00344	rs983494	0.001565	49	0.00344
PAPOLG	2	60836886	16	54.764	0.01058	rs4672410	0.0009812	80	0.00032	16	50.14943	0.01714	rs1432295	0.003367	82	0.000471
REL	2	60962255	16	112.4797	0.00032	rs12713430	0.0002975	80	0.00032	16	107.5495	0.000471	rs13031237	4.35E-05	82	0.000471
PUS10	2	61022607	16	101.1383	0.000445	rs1177279	0.0002473	80	0.00032	16	99.31703	0.000511	rs13031237	4.35E-05	82	0.000471
PEX13	2	61098373	10	64.32374	0.001132	rs1177279	0.0002473	80	0.00032	10	30.85147	0.03566	rs12988616	0.000633	82	0.000471
KIAA1841	2	61146509	14	78.35492	0.001928	rs1177279	0.0002473	80	0.00032	14	45.37361	0.02417	rs12988616	0.000633	82	0.000471
AHSA2	2	61258324	8	22.29219	0.03629	rs2290324	0.004378	80	0.00032	NA	NA	NA	NA	NA	NA	NA
USP34	2	61268093	20	54.96053	0.03312	rs2290324	0.004378	80	0.00032	NA	NA	NA	NA	NA	NA	NA
XPO1	2	61558572	9	27.1636	0.03235	rs778139	0.01006	80	0.00032	NA	NA	NA	NA	NA	NA	NA
NPAL1	4	47713547	12	36.36775	0.02011	rs4340770	0.002227	173	0.000171	NA	NA	NA	NA	NA	NA	NA
TXK	4	47763166	19	116.4092	0.000171	rs17471024	6.21E-05	173	0.000171	20	109.7325	0.000392	rs17470892	1.44E-05	153	0.000392
TEC	4	47832556	33	137.1188	0.000717	rs17471024	6.21E-05	173	0.000171	33	142.8695	0.000476	rs17470892	1.44E-05	153	0.000392
IL3	5	1.31E+08	11	38.55111	0.01816	rs3864277	0.00077	206	0.01117	11	39.55668	0.01759	rs152198	0.002281	195	0.000809
CSF2	5	1.31E+08	10	43.20431	0.01117	rs3864277	0.00077	206	0.01117	10	47.58008	0.00755	rs152198	0.002281	195	0.000809

P4HA2	5	1.32E+08	13	40.41141	0.02531	rs715285	0.00397	206	0.01117	13	74.17733	0.00087	rs2077380	3.34E-05	195	0.000809
PDLIM4	5	1.32E+08	NA	NA	NA	NA	NA	NA	NA	18	95.06358	0.000809	rs2077380	3.34E-05	195	0.000809
SLC22A4	5	1.32E+08	NA	NA	NA	NA	NA	NA	NA	22	75.75723	0.0115	rs3792884	6.20E-05	195	0.000809
LOC441108	5	1.32E+08	NA	NA	NA	NA	NA	NA	NA	23	78.76226	0.00855	rs2070729	0.004761	196	0.00855
IRF1	5	1.32E+08	16	66.37171	0.00359	rs4143832	0.00172	207	0.000518	16	54.99436	0.01076	rs2070729	0.004761	196	0.00855
IL5	5	1.32E+08	6	37.62761	0.000518	rs4143832	0.00172	207	0.000518	NA	NA	NA	NA	NA	NA	NA
RAD50	5	1.32E+08	9	33.87905	0.0191	rs4143832	0.00172	207	0.000518	NA	NA	NA	NA	NA	NA	NA
EIF3B	7	2360999	11	75.60162	5.20E-05	rs6952809	3.37E-05	266	0.000052	11	28.02535	0.03787	rs1548597	0.002413	257	0.03787
CHST12	7	2409784	13	66.20979	0.000148	rs6952809	3.37E-05	266	0.000052	13	29.65083	0.04015	rs1548597	0.002413	257	0.03787
RCL1	9	4782833	NA	NA	NA	NA	NA	NA	NA	34	69.27655	0.04209	rs7038267	0.0203	306	0.01543
JAK2	9	4975244	12	40.51604	0.00996	rs12347727	0.00724	322	0.00996	12	36.5556	0.01543	rs17718680	0.000531	306	0.01543
INSL6	9	5153862	6	37.92267	0.01117	rs10491651	0.0089	322	0.00996	NA	NA	NA	NA	NA	NA	NA
INSL4	9	5221418	5	31.9956	0.01026	rs10491651	0.0089	322	0.00996	NA	NA	NA	NA	NA	NA	NA
PAX5	9	36828530	31	65.82323	0.02318	rs7020413	0.00813	331	0.02318	29	54.7537	0.04573	rs3824344	0.004943	315	0.04573
SCGB1C1	11	183079	16	60.50627	0.01027	rs4980335	0.00104	374	0.000924	NA	NA	NA	NA	NA	NA	NA
ODF3	11	186760	19	67.05872	0.00986	rs4980335	0.00104	374	0.000924	19	51.41517	0.03122	rs498217	0.008865	364	0.0092
BET1L	11	192923	22	77.35674	0.00635	rs4980335	0.00104	374	0.000924	22	58.84193	0.02568	rs498217	0.008865	364	0.0092
RIC8A	11	198529	23	80.34385	0.00608	rs4980335	0.00104	374	0.000924	23	59.7456	0.02786	rs498217	0.008865	364	0.0092
SIRT3	11	205029	25	83.58659	0.00665	rs4980335	0.00104	374	0.000924	25	65.37202	0.02608	rs498217	0.008865	364	0.0092
PSMD13	11	226807	26	89.1471	0.00537	rs4980335	0.00104	374	0.000924	26	69.14908	0.02084	rs498217	0.008865	364	0.0092
NLRP6	11	268569	18	72.15861	0.001152	rs10398	0.00011	374	0.000924	18	53.55222	0.0092	rs498217	0.008865	364	0.0092
ATHL1	11	279137	16	57.02261	0.00169	rs10398	0.00011	374	0.000924	16	41.89066	0.01512	rs498217	0.008865	364	0.0092
IFITM5	11	288202	12	43.73017	0.00249	rs10398	0.00011	374	0.000924	12	29.51358	0.02492	rs498217	0.008865	364	0.0092
IFITM2	11	298106	10	32.64602	0.002805	rs10398	0.00011	374	0.000924	NA	NA	NA	NA	NA	NA	NA
IFITM1	11	303990	10	32.64602	0.00308	rs10398	0.00011	374	0.000924	NA	NA	NA	NA	NA	NA	NA
IFITM3	11	309672	8	29.01861	0.000924	rs10398	0.00011	374	0.000924	NA	NA	NA	NA	NA	NA	NA
F2	11	46697318	NA	NA	NA	NA	NA	NA	NA	6	19.21001	0.02747	rs4752932	0.009175	376	0.000912
CKAP5	11	46721659	NA	NA	NA	NA	NA	NA	NA	7	32.09703	0.01242	rs2306033	0.007294	376	0.000912
LRP4	11	46834993	NA	NA	NA	NA	NA	NA	NA	6	46.58391	0.000912	rs11039035	0.000869	376	0.000912

C11orf49	11	46914826	NA	NA	NA	NA	NA	NA	NA	17	97.64172	0.00258	rs11039035	0.000869	376	0.000912
ARFGAP2	11	47142427	NA	NA	NA	NA	NA	NA	NA	11	32.17387	0.0311	rs4752965	0.001204	376	0.000912
DDB2	11	47193068	10	46.32739	0.00125	rs7120118	0.00119	392	0.000693	12	41.07689	0.01674	rs4752965	0.001204	376	0.000912
ACP2	11	47217428	10	54.34377	0.000699	rs7120118	0.00119	392	0.000693	10	45.11937	0.001639	rs7120118	0.002127	376	0.000912
ACP2	11	47217428	10	54.34377	0.000727	rs7120118	0.00119	392	0.000693	10	40.17069	0.00449	rs7120118	0.002127	376	0.000912
NR1H3	11	47227024	10	54.34377	0.000693	rs7120118	0.00119	392	0.000693	10	40.17069	0.00472	rs7120118	0.002127	376	0.000912
MADD	11	47247534	11	63.66385	0.000801	rs7120118	0.00119	392	0.000693	12	48.97574	0.00493	rs10838692	0.001829	376	0.000912
MYBPC3	11	47309532	8	32.96426	0.00924	rs10838692	0.00122	392	0.000693	9	30.67595	0.01528	rs10838692	0.001829	376	0.000912
NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	11	30.85144	0.04229	rs10838692	0.001829	376	0.000912
CCDC85B	11	65414450	7	18.73915	0.04959	rs500161	0.002054	399	0.000038	NA	NA	NA	NA	NA	NA	NA
FOSL1	11	65416267	8	35.90855	0.00701	rs606978	3.42E-05	399	0.000038	8	27.04747	0.02244	rs500161	0.003384	382	0.001148
C11orf68	11	65440858	7	50.11993	0.000614	rs606978	3.42E-05	399	0.000038	8	33.0662	0.01356	rs500161	0.003384	382	0.001148
DRAP1	11	65443303	7	50.11993	0.000649	rs606978	3.42E-05	399	0.000038	8	33.0662	0.01268	rs500161	0.003384	382	0.001148
TSGA10IP	11	65469690	12	90.65737	0.000269	rs606978	3.42E-05	399	0.000038	13	83.55566	0.001171	rs754532	0.000103	382	0.001148
SART1	11	65485735	11	93.4897	0.000112	rs606978	3.42E-05	399	0.000038	12	76.73825	0.001148	rs754532	0.000103	382	0.001148
EIF1AD	11	65520592	11	78.72116	9.50E-05	rs613924	5.81E-05	399	0.000038	12	60.73457	0.001775	rs754532	0.000103	382	0.001148
BANF1	11	65526125	12	91.65904	3.80E-05	rs613924	5.81E-05	399	0.000038	13	60.87515	0.00254	rs754532	0.000103	382	0.001148
CST6	11	65536037	11	76.01615	7.50E-05	rs613924	5.81E-05	399	0.000038	12	57.10035	0.001942	rs754532	0.000103	382	0.001148
CATSPER1	11	65540798	11	76.01615	7.20E-05	rs613924	5.81E-05	399	0.000038	11	53.21847	0.00157	rs754532	0.000103	382	0.001148
GAL3ST3	11	65564811	10	66.21519	0.000157	rs613924	5.81E-05	399	0.000038	10	24.98483	0.04717	rs747526	0.0003	382	0.001148
SF3B2	11	65576391	9	51.36033	0.00114	rs503156	0.000322	399	0.000038	NA	NA	NA	NA	NA	NA	NA
PACS1	11	65594399	17	50.18067	0.02364	rs503156	0.000322	399	0.000038	NA	NA	NA	NA	NA	NA	NA
ETS1	11	1.28E+08	34	81.39607	0.01218	rs4245079	0.00044	418	0.01218	34	75.25326	0.02008	rs7122859	0.001589	395	0.02008
BCL7A	12	1.21E+08	5	13.34707	0.04572	rs11043307	0.04298	451	0.000919	5	35.92612	0.000254	rs11835818	0.000236	428	0.000046
MLXIP	12	1.21E+08	14	73.20462	0.00355	rs2292443	0.0003	451	0.000919	14	106.1202	0.000163	rs1047796	0.00013	428	0.000046
LRRC43	12	1.21E+08	13	64.22273	0.00276	rs2292443	0.0003	451	0.000919	13	79.93842	0.000357	rs1047796	0.00013	428	0.000046
IL31	12	1.21E+08	12	63.44227	0.00239	rs2292443	0.0003	451	0.000919	12	72.28707	0.000677	rs1047796	0.00013	428	0.000046
B3GNT4	12	1.21E+08	8	45.69593	0.001238	rs2292443	0.0003	451	0.000919	9	40.10194	0.00448	rs11611988	0.003876	428	0.000046
DIABLO	12	1.21E+08	8	44.71881	0.000919	rs2292443	0.0003	451	0.000919	9	46.97952	0.001075	rs907482	0.00203	428	0.000046

VPS33A	12	1.21E+08	7	27.38524	0.00909	rs2292443	0.0003	451	0.000919	8	43.04591	0.001237	rs4758670	0.000971	428	0.000046
CLIP1	12	1.21E+08	11	45.92899	0.00495	rs7301515	0.00297	451	0.000919	10	72.45726	4.60E-05	rs10846921	0.000313	428	0.000046
ZCCHC8	12	1.22E+08	9	33.44878	0.02227	rs7301515	0.00297	451	0.000919	8	64.37445	0.000403	rs10846921	0.000313	428	0.000046
RSRC2	12	1.22E+08	9	28.58497	0.03052	rs7301515	0.00297	451	0.000919	9	65.89722	0.00045	rs10846921	0.000313	428	0.000046
RSRC2	12	1.22E+08	NA	NA	NA	NA	NA	NA	NA	9	65.89722	0.000439	rs10846921	0.000313	428	0.000046
KNTC1	12	1.22E+08	NA	NA	NA	NA	NA	NA	NA	14	77.3776	0.001221	rs7968222	0.000848	428	0.000046
RCOR1	14	1.02E+08	NA	NA	NA	NA	NA	NA	NA	19	74.73825	0.00376	rs8010152	0.000166	465	0.00113
TRAF3	14	1.02E+08	21	103.0882	0.001292	rs12148050	3.48E-05	487	0.001292	21	104.1795	0.00113	rs941726	7.51E-05	465	0.00113
AMN	14	1.02E+08	8	33.66006	0.01123	rs2273393	0.01305	487	0.001292	NA	NA	NA	NA	NA	NA	NA
CDC42BPB	14	1.02E+08	16	41.81444	0.02619	rs2273393	0.01305	487	0.001292	NA	NA	NA	NA	NA	NA	NA
MARK3	14	1.03E+08	21	86.42348	0.00663	rs3783394	0.00076	488	0.00663	21	154.7726	9.60E-05	rs1467561	1.69E-05	466	0.000096
CKB	14	1.03E+08	13	43.58392	0.02412	rs1467561	0.00218	488	0.00663	13	91.9596	0.000352	rs1467561	1.69E-05	466	0.000096
C14orf172	14	1.03E+08	11	32.02248	0.04165	rs1467561	0.00218	488	0.00663	11	67.16573	0.00136	rs1467561	1.69E-05	466	0.000096
BAG5	14	1.03E+08	NA	NA	NA	NA	NA	NA	NA	10	39.16304	0.01533	rs1467561	1.69E-05	466	0.000096
C14orf153	14	1.03E+08	NA	NA	NA	NA	NA	NA	NA	12	49.38119	0.01032	rs1467561	1.69E-05	466	0.000096
KLC1	14	1.03E+08	NA	NA	NA	NA	NA	NA	NA	14	44.81477	0.02751	rs11160755	0.02118	466	0.000096
XRCC3	14	1.03E+08	NA	NA	NA	NA	NA	NA	NA	10	26.30876	0.04738	rs8007903	0.02955	466	0.000096
ZFYVE21	14	1.03E+08	NA	NA	NA	NA	NA	NA	NA	10	26.30876	0.04612	rs8007903	0.02955	466	0.000096
ITGAM	16	31178788	13	56.1471	0.003134	rs7193268	0.00089	514	0.002478	13	35.52077	0.03216	rs1143678	0.00924	492	0.01925
ITGAX	16	31274009	12	50.7809	0.002478	rs7193268	0.00089	514	0.002478	12	35.4042	0.01925	rs1143678	0.00924	492	0.01925
ITGAD	16	31312133	12	32.05143	0.02253	rs11574637	0.00325	514	0.002478	NA	NA	NA	NA	NA	NA	NA
C17orf57	17	42756348	12	40.98834	0.01166	rs4968318	0.003165	548	0.000157	12	53.79474	0.00259	rs9893901	3.03E-05	530	0.0000001
NPEPPS	17	42963442	4	36.90127	0.000519	rs11079784	0.0001316	548	0.000157	4	73.46174	1E-07	rs11079784	4.36E-07	530	0.0000001
KPNB1	17	43082273	7	62.55541	0.000157	rs11079784	0.0001316	548	0.000157	7	105.8437	1E-07	rs11079784	4.36E-07	530	0.0000001
TBKBP1	17	43127628	8	50.71565	0.000898	rs16942199	0.0001784	548	0.000157	8	90.2871	5.00E-06	rs8078686	6.48E-07	530	0.0000001
TBX21	17	43165608	8	33.93278	0.00371	rs16942199	0.0001784	548	0.000157	8	50.93979	2.00E-04	rs4794053	4.95E-06	530	0.0000001
UNK	17	71292275	NA	NA	NA	NA	NA	NA	NA	10	34.25729	0.00895	rs3744017	0.00012	540	0.000681
UNC13D	17	71334901	NA	NA	NA	NA	NA	NA	NA	7	43.89598	0.000681	rs3744017	0.00012	540	0.000681
WBP2	17	71353374	7	22.24762	0.02328	rs17581728	0.00571	560	0.00105	7	39.17802	0.001208	rs3744017	0.00012	540	0.000681

TRIM47	17	71381839	10	28.52824	0.03587	rs17581728	0.00571	560	0.00105	10	55.4887	0.00205	rs3744017	0.00012	540	0.000681
TRIM65	17	71396635	10	36.8546	0.02248	rs1135889	0.00182	560	0.00105	10	56.14763	0.00417	rs3744017	0.00012	540	0.000681
MRPL38	17	71406318	10	36.8546	0.02192	rs1135889	0.00182	560	0.00105	10	56.14763	0.0038	rs3744017	0.00012	540	0.000681
FBF1	17	71418212	13	38.367	0.02324	rs1135889	0.00182	560	0.00105	13	51.43853	0.00657	rs3744017	0.00012	540	0.000681
ACOX1	17	71449186	15	48.7562	0.00969	rs2069528	0.00162	560	0.00105	NA	NA	NA	NA	NA	NA	NA
CDK3	17	71508581	9	22.59152	0.02884	rs2069528	0.00162	560	0.00105	NA	NA	NA	NA	NA	NA	NA
EVPL	17	71514521	11	31.04031	0.01304	rs2069528	0.00162	560	0.00105	NA	NA	NA	NA	NA	NA	NA
SRP68	17	71546785	10	51.91117	0.00105	rs2305347	0.00044	560	0.00105	NA	NA	NA	NA	NA	NA	NA
GALR2	17	71582486	7	39.19201	0.00272	rs2305347	0.00044	560	0.00105	NA	NA	NA	NA	NA	NA	NA
ZACN	17	71586903	7	39.19201	0.00265	rs2305347	0.00044	560	0.00105	NA	NA	NA	NA	NA	NA	NA
EXOC7	17	71588681	9	43.08315	0.00267	rs2305347	0.00044	560	0.00105	NA	NA	NA	NA	NA	NA	NA
KLF2	19	16296650	8	36.62655	6.00E-04	rs6512102	0.0003152	590	0.000311	8	31.92793	0.00173	rs6512102	1.26E-06	559	0.0000001
EPS15L1	19	16333407	16	88.48963	0.000311	rs11878602	0.0002781	590	0.000311	16	180.1105	1E-07	rs11878602	2.37E-07	559	0.0000001
CALR3	19	16450874	7	36.32593	0.002296	rs11878602	0.0002781	590	0.000311	7	100.8443	1E-07	rs11878602	2.37E-07	559	0.0000001
C19orf44	19	16468204	7	30.63249	0.00592	rs1544768	0.001654	590	0.000311	7	78.28756	6.00E-06	rs10412060	2.06E-06	559	0.0000001
CHERP	19	16489699	9	50.4809	0.00176	rs754292	0.001406	590	0.000311	9	94.55979	1.60E-05	rs10412060	2.06E-06	559	0.0000001
SLC35E1	19	16521647	10	48.91913	0.00242	rs754292	0.001406	590	0.000311	10	79.88877	5.50E-05	rs11667601	7.23E-05	559	0.0000001
MED26	19	16546717	12	53.63174	0.000963	rs754292	0.001406	590	0.000311	12	65.92748	0.000201	rs11667601	7.23E-05	559	0.0000001
C19orf42	19	16617958	NA	NA	NA	NA	NA	NA	NA	8	26.50736	0.00741	rs3786601	7.96E-05	559	0.0000001
TMEM38A	19	16632937	NA	NA	NA	NA	NA	NA	NA	8	19.84108	0.03002	rs3786601	7.96E-05	559	0.0000001
IL12RB1	19	18031370	14	78.26269	0.000143	rs740691	8.83E-07	592	0.0000001	NA	NA	NA	NA	NA	NA	NA
MAST3	19	18069602	15	155.0089	1.00E-07	rs874628	4.32E-08	592	0.0000001	15	46.74158	0.01644	rs874628	0.00228	560	0.00872
PIK3R2	19	18125015	16	181.4169	1.00E-07	rs874628	4.32E-08	592	0.0000001	16	52.53453	0.01093	rs4808762	0.000788	560	0.00872
IFI30	19	18145578	15	151.5595	1.00E-06	rs874628	4.32E-08	592	0.0000001	15	50.68946	0.00872	rs4808762	0.000788	560	0.00872
FKSG24	19	18165039	15	136.032	1.00E-06	rs874628	4.32E-08	592	0.0000001	15	49.3563	0.01059	rs4808762	0.000788	560	0.00872
RAB3A	19	18168610	14	112.6513	1.10E-05	rs874628	4.32E-08	592	0.0000001	14	44.74111	0.01381	rs4808762	0.000788	560	0.00872
PDE4C	19	18179770	21	194.0875	3.00E-06	rs874628	4.32E-08	592	0.0000001	22	54.50826	0.03916	rs4808762	0.000788	560	0.00872
KIAA1683	19	18228907	19	131.659	2.20E-05	rs12608504	6.53E-07	592	0.0000001	NA	NA	NA	NA	NA	NA	NA
JUND	19	18251562	15	90.5547	0.000287	rs12608504	6.53E-07	592	0.0000001	NA	NA	NA	NA	NA	NA	NA

LSM4	19	18278716	21	93.28058	0.000142	rs12608504	6.53E-07	592	0.0000001	NA	NA	NA	NA	NA	NA	NA
MN1	22	26474264	7	19.47515	0.01412	rs9625312	0.00155	651	0.00592	NA	NA	NA	NA	NA	NA	NA
PITPNB	22	26577656	8	30.59403	0.00592	rs12169097	0.00449	651	0.00592	8	24.8884	0.01738	rs12170161	0.006413	606	0.01738
PVALB	22	35526690	11	42.14601	0.00525	rs4821544	0.0001105	656	0.000001	10	36.73486	0.00438	rs741997	0.000261	607	0.00438
NCF4	22	35586975	18	174.2993	1.00E-06	rs2413436	8.72E-06	656	0.000001	18	56.4745	0.01512	rs741997	0.000261	607	0.00438
CSF2RB*	22	35639620	18	139.1452	2.80E-05	rs2413436	8.72E-06	656	0.000001	18	27.62367	0.163	rs2413436	0.01541	NA	0.163

* CSF2RB is not significant in meta2.5 (p=0.163)

rows in red represent genes with p-values in both studies and identified in the intersection network (candidates)

Authors & affiliations

Authors:
Sergio E. Baranzini[1], Pouya Khankhanian[1], Nikolaos A. Patsopoulos[2,45], Michael Li[1], Jim Stankovich[3], Chris Cotsapas[4,46,48], Helle Bach Søndergaard[5], Maria Ban[6], Nadia Barizzone[7], Laura Bergamaschi[7], David Booth[8], Dorothea Buck[9], Paola Cavalla[10], Elisabeth G. Celius[11], Manuel Comabella[12], Giancarlo Comi[13], Alastair Compston[6], Isabelle Cournu-Rebeix[14], Sandra D'alfonso[7], Vincent Damotte[13], Lennox Din[1], Bénédicte Dubois[15], Irina Elovaara[15], Federica Esposito[32], Bertrand Fontaine[14], Andre Franke[17], An Goris[15], Pierre-Antoine Gourraud[1], Christiane Graetz[18], Franca R. Guerini[19], Léna Guillot-Noel[14], David Hafler[4], Hakon Hakonarson[20], Per Hall[21], Anders Hamsten[22], Hanne F. Harbo[23], Bernhard Hemmer[24], Jan Hillert[25], Anu Kemppinen[6], Ingrid Kockum[25], Keijo Koivisto[26], Malin Larsson[27], Mark Lathrop[28], Maurizio Leone[29], Christina M. Lill[18], Fabio Macciardi[30], Roland Martin[31], Vittorio Martinelli[13], Filippo Martinelli-Boneschi[32], Jacob L. McCauley[33], Kjell-Morten Myhr[34], Paola Naldi[29], Tomas Olsson[25], Annette Oturai[5], Margaret A. Pericak-Vance[33], Franco Perla[35], Mauri Reunanen[36], Janna Saarela[37], Safa Saker-Delye[38], Marco Salvetti[39], Finn Sellebjerg[5], Per Soelberg Sørensen[5], Anne Spurkland[40], Graeme Stewart[8], Bruce Taylor[41], Pentti Tienari[42], Juliane Winkelmann[9,48,50], WTCCC2[*], Frauke Zipp[18], Adrian J. Ivinson[43], Jonathan L. Haines[44], Stephen Sawcer[6], Philip DeJager[45], Stephen L. Hauser[1], Jorge R. Oksenberg[1]

Affiliations:

1 Department of Neurology, University of California at San Francisco, Sandler Neurosciences Center, 675 Nelson Rising Lane, San Francisco, CA 94148, USA
2 Program in Translational NeuroPsychiatric Genomics, Institute for the Neurosciences, Department of Neurology, Brigham & Women's Hospital, Boston, MA, USA
3 Menzies Research Institute Tasmania, University of Tasmania, Private Bag 23, Hobart, Tasmania 7000, Australia.
4 Program in Medical and Population Genetics, Broad Institute of Harvard University and MIT, Cambridge, 02142 MA, USA
5 Danish Multiple Sclerosis Center_ Department of Neurology_ Copenhagen University Hospital_ Rigshospitalet_ section 2082_ 2100 Copenhagen_ Denmark
6 University of Cambridge, Department of Clinical Neurosciences, Addenbrookes Hospital, BOX 165, Hills Road, Cambridge CB2 0QQ, UK
7 Department of Health Sciences and Interdisciplinary Research Center of Autoimmune Diseases (IRCAD), University of Eastern Piedmont, Novara, Italy
8 Westmead Millennium Institute, University of Sydney, New South Wales, Australia
9 Department of Neurology, Klinikum Rechts der Isar, Technische Universität München, Ismaninger Strasse 22, 81675 Munich, Germany.
10 Department of Neurology, Ospedale San Giovanni Battista,Torino, Italy
11 Department of Neurology_ Oslo University Hospital_ Ullevål_ Oslo_ Norway

[12] Clinical Neuroinmunology Unit, Multiple Sclerosis Center of Catalonia (CEM-Cat), Vall d'Hebron University Hospital, 08035 Barcelona, Spain
[13] Department of Neurology, INSPE, Division of Neuroscience, San Raffaele Scientific Institute, Via Olgettina 58, 20132, Milan, Italy
[14] INSERM UMR S 975 CRICM, UPMC, Département de neurologie Pitié-Salpêtrière, AP-HP 75013 PARIS, France
[15] Laboratory for Neuroimmunology, Section of Experimental Neurology, KU Leuven, Herestraat 49, 3000 Leuven, Belgium
[16] Department of Neurology_ Tampere University Hospital_ Tampere_ Finland
[17] Institute of Clinical Molecular Biology, Christian-Albrechts-University of Kiel, D-24105, Germany
[18] Focus Program Translational Neuroscience (FTN), Rhine Main Neuroscience Network (rmn2), Johannes Gutenberg University-Medical Center, Mainz, Germany
[19] Don C. Gnocchi Foundation, Milan, Italy
[20] Center for Applied Genomics, The Children's Hospital of Philadelphia, 3615 Civic Center Blvd., Philadelphia, PA 19104, USA
[21] Department of Medical Epidemiology and Biostatistics, Karolinska Institute, 17177 Stockholm, Sweden
[22] Atherosclerosis Research Unit, Department of Medicine Solna, Karolinska Institutet, Center for Molecular Medicine, L8:03, Karolinska University Hospital Solna, S-171 76 Stockholm, Sweden
[23] Department of Neurology, Oslo University Hospital, Ullevål, and University of Oslo, Norway
[24] Department of Neurology, Klinikum rechts der Isar, Technische Universität München, Ismaninger Strasse 22, 81675 Munich, Germany
[25] Department of Clinical Neuroscience, Karolinska Institutet, Stockholm, Sweden
[26] Department of Neurology_ Seinäjoki Central Hospital_ 60220 Seinäjoki_ Finland
[27] IFM Bioinformatics, Linköping University, SE-58183 Linköping, Sweden
[28] Commissariat à l'énergie atomique, Institut de Génomique, Centre National de Genotypage, 2 rue Gaston Cremieux, CP 5721, 91057 Evry Cedex, France
[29] Department of Neurology, Ospedale Maggiore, Novara, Italy
[30] Department of Psychiatry and Human Behavior, University of California at Irvine, 312 Sprague Hall, 92617 Irvine, USA
[31] Institute for Neuroimmunology and Clinical MS Research (inims), Center for Molecular Neurobiology (ZMNH), University Medical Center Hamburg-Eppendorf, Falkenried 94, 20251 Hamburg, Germany
[32] Laboratory of Genetics of Neurological complex disorder, Institute of Experimental Neurology (INSPE), Division of Neuroscience, San Raffaele Scientific Institute, Via Olgettina 58, 20132, Milan, Italy
[33] John P. Hussman Institute for Human Genomics, University of Miami, Miller School of Medicine, Miami, FL 33136, USA
[34] The Norwegian multiple sclerosis registry and biobank_ Department of Neurology_ Haukeland University Hospital_ Bergen_ Norway
[35] SC Neurologia, Ospedale di Mondovì, Italy

[36] Department of Neurology_ Oulu University Hospital_ 90014 Oulu_ Finland
[37] Institute for Molecular Medicine Finland, University of Helsinki, Helsinki, Finland
[38] Généthon, Banque d'ADN et de cellules, Evry, France
[39] Neurology and Center for Experimental Neurological Therapy (CENTERS), Università La Sapienza, Roma, Italy
[40] Institute of Basic Medical Sciences, University of Oslo, Oslo, Norway
[41] Menzies Research Institute Tasmania, University of Tasmania
[42] Department of Neurology_ Helsinki University Central Hospital and Molecular Neurology Programme_ Biomedicum_ University of Helsinki_ Helsinki_ Finland
[43] Harvard NeuroDiscovery Center, Harvard Medical School, Boston, MA 02115, USA
[44] Center for Human Genetics Research, Vanderbilt University Medical Center, 525 Light Hall, 2215 Garland Avenue, Nashville, TN 37232, USA
[45] Program in Translational NeuroPsychiatric Genomics, Institute for the neurosciences, Departments of Neurology and Psychiatry, Brigham & Women's Hospital, Boston, 02115 MA, USA
[46] Harvard Medical School, Boston, MA, USA
[47] Dept of Genetics, Yale School of Medicine, New Haven, CT 06520, USA
[48] Institute of Human Genetics, Helmholtz Zentrum München, Ingolstädter Stasse 1 München, 85764 Oberschleißheim, Germany
[49] Broad Institute of Harvard and Massachusetts Institute of Technology, Cambridge, MA, USA
Munich Center for Systems Neurology (SyNergy)

Membership of Wellcome Trust Case Control Consortium 2

Management Committee

Peter Donnelly (Chair)[1,2], Ines Barroso (Deputy Chair)[3], Jenefer M Blackwell[4,5], Elvira Bramon[6], Matthew A Brown[7], Juan P Casas[8], Aiden Corvin[9], Panos Deloukas[3], Audrey Duncanson[10], Janusz Jankowski[11], Hugh S Markus[12], Christopher G Mathew[13], Colin NA Palmer[14], Robert Plomin[15], Anna Rautanen[1], Stephen J Sawcer[16], Richard C Trembath[13], Ananth C Viswanathan[17], Nicholas W Wood[18]

Data and Analysis Group

Chris C A Spencer[1], Gavin Band[1], Céline Bellenguez[1], Colin Freeman[1], Garrett Hellenthal[1], Eleni Giannoulatou[1], Matti Pirinen[1], Richard Pearson[1], Amy Strange[1], Zhan Su[1], Damjan Vukcevic[1], Peter Donnelly[1,2]

DNA, Genotyping, Data QC and Informatics Group

Cordelia Langford[3], Sarah E Hunt[3], Sarah Edkins[3], Rhian Gwilliam[3], Hannah Blackburn[3], Suzannah J Bumpstead[3], Serge Dronov[3], Matthew Gillman[3], Emma Gray[3], Naomi Hammond[3], Alagurevathi Jayakumar[3], Owen T McCann[3], Jennifer Liddle[3], Simon C Potter[3], Radhi Ravindrarajah[3], Michelle Ricketts[3], Matthew Waller[3], Paul Weston3, Sara Widaa[3], Pamela Whittaker[3], Ines Barroso3, Panos Deloukas[3].

Publications Committee

Christopher G Mathew (Chair)[13], Jenefer M Blackwell[4,5], Matthew A Brown[7], Aiden Corvin[9], Chris C A Spencer[1]

[1] Wellcome Trust Centre for Human Genetics, University of Oxford, Roosevelt Drive, Oxford OX3 7BN, UK
[2] Dept Statistics, University of Oxford, Oxford OX1 3TG, UK
[3] Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, UK
[4] Telethon Institute for Child Health Research, Centre for Child Health Research, University of Western Australia, 100 Roberts Road, Subiaco, Western Australia 6008
[5] Cambridge Institute for Medical Research, University of Cambridge School of Clinical Medicine, Cambridge CB2 0XY, UK
[6] Department of Psychosis Studies, NIHR Biomedical Research Centre for Mental Health at the Institute of Psychiatry, King's College London and The South London and Maudsley NHS Foundation Trust, Denmark Hill, London SE5 8AF, UK
[7] University of Queensland Diamantina Institute, Brisbane, Queensland, Australia
[8] Dept Epidemiology and Population Health, London School of Hygiene and Tropical Medicine, London WC1E 7HT and Dept Epidemiology and Public Health, University College London WC1E 6BT, UK
[9] Neuropsychiatric Genetics Research Group, Institute of Molecular Medicine, Trinity College Dublin, Dublin 2, Eire
[10] Molecular and Physiological Sciences, The Wellcome Trust, London NW1 2BE
[11] Department of Oncology, Old Road Campus, University of Oxford, Oxford OX3 7DQ, UK , Digestive Diseases Centre, Leicester Royal Infirmary, Leicester LE7 7HH, UK and Centre for Digestive Diseases, Queen Mary University of London, London E1 2AD, UK
[12] Clinical Neurosciences, St George's University of London, London SW17 0RE
[13] King's College London Dept Medical and Molecular Genetics, King's Health Partners, Guy's Hospital, London SE1 9RT, UK
[14] Biomedical Research Centre, Ninewells Hospital and Medical School, Dundee DD1 9SY, UK
[15] King's College London Social, Genetic and Developmental Psychiatry Centre, Institute of Psychiatry, Denmark Hill, London SE5 8AF, UK
[16] University of Cambridge Dept Clinical Neurosciences, Addenbrooke's Hospital, Cambridge CB2 0QQ, UK
[17] NIHR Biomedical Research Centre for Ophthalmology, Moorfields Eye Hospital NHS Foundation Trust and UCL Institute of Ophthalmology, London EC1V 2PD, UK
[18] Dept Molecular Neuroscience, Institute of Neurology, Queen Square, London WC1N 3BG, UK.

Acknowledgements (continuation)

This project was supported by the National Institutes of Health (NS049477, NS26799), the US National Multiple Sclerosis Society (grant RG 4198-A-1), the Australian National Health and Medical Research Council (NHMRC), the Australian Research Council Linkage Program Grant, the JHH Charitable Trust Fund, Multiple Sclerosis Research Australia, the Health Research Council New Zealand, the National MS Society of New Zealand, Wetenschappelijk Onderzoek Multiple Sclerose, the Bayer Chair on fundamental genetic research regarding the neuroimmunological aspects of multiple sclerosis, the Belgian Neurological Society, the Cambridge NIHR Biomedical Research Centre, the Danish Multiple Sclerosis Society, the Danish Council for Strategic Research, the Neuropromise EU grant (LSHM-CT-2005-018637), the Center of Excellence for Disease Genetics of the Academy of Finland, the Sigrid Juselius Foundation, Helsinki University Central Hospital Research Foundation, the Bundesministerium für Bildung und Technologie (KKNMS consortium Control MS), Institut National de la Santé et de la Recherche Médicale (INSERM), Fondation d'Aide pour la Recherche sur la Sclérose en Plaques (ARSEP), Association Française contre les Myopathies (AFM), GIS-IBISA,the Institut du Cerveau et de la Moelle épinière (ICM) and the program "Investissements d'avenir" ANR-10-IAIHU-06, the Italian Foundation for Multiple Sclerosis (FISM grants, 2008/R/11 and 2008/R/15, "Progetto Speciale Immunochip", 2011/R/14), Fondazione Cariplo (grant 2010-0728), the Italian Ministry of Health (grant Giovani Ricercatori 2007 - D.lgs 502/92), Regione Piemonte (grants 2003, 2004, 2008, 2009), CRT Foundation, Turin, the Norwegian MS Register and Biobank, The Research Council of Norway, South-Eastern and Western Norway regional Health Authories, Ullevål University Hospital Scientific Advisory Council, Haukeland University Hospital, the Amici Centro Sclerosi Multipla del San Raffaele, the Association of British Neurologists, Spanish Ministry of Health(FISPI060117), Bibbi and Niels Jensens Foundation, Montel Williams foundation, Hjärnfonden and the Swedish medical research council, the Swedish Council for Working life and Social Research, the Gemeinnützige Hertie Stiftung, Northern California Kaiser Permanente members and the Polpharma Foundation.

We acknowledge use of data generated during the Wellcome Trust Case Control Consortium 2 project (funded by the Wellcome Trust - 085475/B/08/Z, 085475/Z/08/Z and 084702/Z/08/Z) which includes data from the British 1958 Birth Cohort DNA collection (funded by the Medical Research Council grant G0000934 and the Wellcome Trust grant 068545/Z/02) and the UK National Blood Service controls (funded by the Wellcome Trust). We also acknowledge the use of external control data from the popgen biobank (supported by the German Ministry of Education and Research (BMBF) through the National Genome Research Network (NGFN) and received infrastructure support through the DFG excellence cluster "Inflammation at Interfaces"), the KORA research platform (KORA: Cooperative Research in the Region of Augsburg) and the MONICA Augsburg studies (funded by the German Federal Ministry of Education, Science, Research and Technology and by the State of Bavaria), the Swedish Breast Cancer study (funded by the Agency for Science & Technology and Research of Singapore

(A*STAR), the Susan G Komen Breast Cancer Foundation, and the National Institute of Health - R01 CA 104021), HYPERGENES (a European Commission project, HEALTH-F4-2007-201550), the Children's Hospital of Philadelphia (CHOP, funded by an Institutional Development Award to the Center for Applied Genomics from the Children's Hospital of Philadelphia), the Swedish CAD study (funded by the Knut and Alice Wallenberg Foundation, the Swedish Research Council - 8691 and the Stockholm County Council - 562183) and the collaboration of the BRC-REFGENSEP, Pitié-Salpêtrière Centre d'Investigation Clinique (CIC) and Généthon.

AJHG Conflict of Interest Form

Cell Press, 600 Technology Square, 5th floor, Cambridge, MA 02139, USA

Please complete this form electronically and upload the file with your final submission.

AJHG requires all authors to disclose any financial interest that might be construed to influence the results or interpretation of their manuscript.

As a guideline, any affiliation associated with a payment or financial benefit exceeding $10,000 p.a. or 5% ownership of a company or research funding by a company with related interests would constitute a financial interest that must be declared. This policy applies to all submitted research manuscripts and review material.

Examples of statement language include: AUTHOR is an employee and shareholder of COMPANY; AUTHOR is a founder of COMPANY and a member of its scientific advisory board. This work was supported in part by a grant from COMPANY.

Please disclose any such interest below on behalf of all authors of this manuscript.

Please check one of the following:

☒ None of the authors of this manuscript have a financial interest related to this work.

☐ Please print the following disclosure statement in the Acknowledgments section:

Please provide the following information:

☒ Please check this box to indicate that you have asked every author of this work to declare any conflicts of interest. Your answers on this form are on behalf of every author of this work.

Manuscript # (research articles only): AJHG-D-13-00106R2

Title: Network-based pathway analysis in multiple sclerosis with GWAS data from 15,000 cases and 30,000 controls

Author list: The International Multiple Sclerosis Genetics Consortium

Your name: Sergio E. Baranzini

Date: 4/22/2013